

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: Wesley - July

Searcher Phone #: 308-4501 ✓

Searcher Location: Beotel Lib.

Date Searcher Picked Up: 6/7/01

Date Completed: 6/7/01

Searcher Prep & Review Time: _____

Clerical Prep Time: 9 min.

Online Time: 2 m

Type of Search

NA Sequence (#)_____

AA Sequence (#) 1

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext

Patent Family _____

Other \$ 1,000.00

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Orbit _____

Dr. Link _____

Lexis/Nexis _____

Sequence Systems AB5302

WWW/Internet _____

Other (specify) _____

THIS PAGE BLANK (USPTO)

FT	Region	/note="WD motif" 516..544 /note="WD motif"
FT		
PN	MO9938969-A1.	
XX		
XX	05-AUG-1999.	
XX		
XX	29-JAN-1999;	99MO-FR00196.
XX		
XX	09-DEC-1998;	98FR-0015545.
XX	30-JAN-1998;	98FR-0001100.
XX		
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE. (INSP) INST PASTEUR.	
PI	Arenzana Seisdedos F, Benarous R, Concordet J, Durand H; Kroll M, Margotlin F;	
PI		
DR	WPI: 1999-469329/39.	
DR	N-PSDB: X86501.	
XX		
PT	New human beta-transducin repeat containing protein and its fragments useful as, or to screen for, antiviral, antitumour, anti-inflammatory and anti-Alzheimer's agents	
XX		
PS	Claim 1; Page 60-61; 71pp; French.	
XX		
CC	The present sequence represents a human beta-transducin repeat containing protein (beta-Trcp). The protein directs proteins to the proteosome degradation pathways. The protein is able to interact with the Vpu protein of human immune deficiency virus-1 (HIV-1), cellular proteins IkappaB or beta-catenin (bc) and/or protein Skp1. The protein controls ubiquitinylation of phosphorylated proteins and thus their targeting to proteosomes for degradation. Depending on whether the process is inhibited or promoted, the result may be delayed breakdown of CD4 (in cases of HIV-1 infection); increased activity of Ikb (and thus reduced activity of NFkappaB) and increased degradation of mutant bc in tumour cells, or increased bc survival (and reduced apoptosis) in Alzheimer's patients. The beta-Trcp protein, and its active peptide fragments, or its nucleic acid, are used to screen for anti HIV-1 agents (antivirals), antitumour agents that disrupt cell cycle regulation or protein degradation in human tumour cells, and anti-inflammatory agents that disrupt activation of NFkappaB. Fragments of the protein are also useful for treating osteo-articular inflammation or acute inflammation associated with release of tumour necrosis factor.	
XX	Sequence	569 AA;

	Query Match	Similarity	Score	DB	Length
	Best Local	100.0%	3034	20	569
	Matches	569	Conservative	0	Mismatches
				0	Indels
				0	Gaps
QY	1	MDPEAVLQEKALFKPNSSEREDCNGNPEPRKIIPEKSLRQTYNSCARCLNDETVCGLA	60		
Db	1	mnpaevavgakalkfmseseredcngpeprkiipeknsliqynsarcilnqetvcia	60		
QY	61	STAKTEENCYAKTKLLANGTSSMIVPQRKRLASAYKEKEKLCVKAFEEOMSESDDQVEFEVHL	120		
Db	61	stankteencyaktklangtssmivpqrkrlasaykekelevkyfeqsesddqvefvehl	120		
QY	121	ISQMCYQHGHINSYLPKMLQDEFTTALPARGLDIIAENILSYLDAKSLCAELVCKEWY	180		
Db	121	isqmcyqhghinsylpkmlqdefttalparjldhiaenilslsydaksllcaaelvckewy	180		
QY	181	RYNSDGLMKKLLIERVWRPDSLMRGLAERGGVTLFNKRPDGNAPNSFYRLVPKII	240		
Db	181	rvsdglmkllierwrtcdslmrjlaerwgsytlfnkppdgnapnsfyrlvypkii	240		
QY	241	ODFTTSSNRCGRHSIORSEFSKGVCLQYDDOKTVSGLRDMTKIMQNTLECK	300		
Db	241	qdfttlesnwrgrhsiqrihrcsefskgyvclqyddoktvsgrldmtkimwntleek	300		

Qy	301	RILGHHGSYLCLOYDEVAITITSSDSSTVRWMDVNTGEMLNTLIHCEAVLHLRFNNGM	360
Dh	301	rlilghgsylclgydervlittgsdsstvrwmdntgelmntllhnceavllhrlfnngm	360
Qy	361	VTCSKDRSIAVMDMASPTDITLRRLVGHRAAVNVDFDDKYIVSASGDRTIKWMTSTC	420
Dh	361	vtcskdrsiavwmdmaspltditlrrvlghraavnvdfddkylvasagdrtlkwmtstc	420
Qy	421	EFVFTLNGHKRGICLQYRDLVLVSSGSSDMTIRLMDTEGCAKRLVEGHELRICIFDN	480
Dh	421	efvftlnghkrgiclqyrdtlvsvsgssdmtrlmdtegcakrlveghelelricifdn	480
Qy	481	KRIVSAYDGKIKYWDLVALDPAPAGTCLRTLVEHSGVRFLQDFEQIVSSHDT	540
Dh	481	kriivsavgdkikwydvlvaldpapagtclrtlvshsgvrflqdfefqivsshdtt	540
Qy	541	ILIMDFLMDPAAOAEPSPSPRTTYTISR	569
Dh	541	ilimdfldmdpaaoaepspspertytysr	569

RESULT	2
B12813	
ID	B12813 standard; protein; 569 AA

AC B12813;

DT 27-NOV-2000 (first entry)

DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.

KM Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;
KM beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; EMD1;
KM gene therapy; colon cancer; beta-transducin repeat containing protein;
KM beta-TTCP.

05 Homo sapiens.

PN JP2000166542-A

PD 20-JUN-2000

PF 02-DEC-1998; 98JP-0343437

PR 02-DEC-1998; 98JP-0343437.

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN

DR WPI; 2000-485550/43.

DR N-PSDB; A73132.

PT F-box protein of ubiquitin ligase SCF complex which promotes the

PS Claim 3; Page 10-12; 19pp; Japanese

The present invention describes an F-box motif/protein of ubiquitin ligase SCF complex which promotes the ubiquitination of IkappaB or beta-actinin and is constituted by Skp1 protein, Cull protein and a complex (SCF complex) of F-box protein containing F-box motif and a repeat motif and has the amino acid sequence of 45 residues (B12811) or one of two 569 residue sequences (B12812, which is mouse ubiquitin ligase FWD1 protein) and (B12813, which is human beta-tubulin repeat containing protein (beta-TrCP)). The F-box protein can be used for the gene therapy of colon cancer by being recombined to a virus vector.

SQ Sequence 569 AA;

Query Match	100.0%;	Score 3034;	DB 21;	Length 569;
Best Local Similarity	100.0%;	Pred. No. 2,4e-287;		
Matches 569; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	MDPAEAVQDEKALRPMNSSEEDCNNGEPKKITIPKNSLRQTYNSCARCLINDQETVCLA	60
Db	1	mdpaevaiqekalkfmsnserecdnngppkkilipeknslrqtynscaarlclndqetvcla	60
QY	61	STAMTEKCVKTKTLKANGTSSMIVPKGRKLSASYEKEKEKLVKFEQWSESDQVEFEHL	120
Db	61	stamtkencvaktklangtssmivpkqrklsasyekekelcvkyfeqwsedqvefehl	120
QY	121	ISOMCHVOHGHSINSLKRPMLQDRFITALPARGLDHIAENILSYLDKASTLCAELVCKEWY	180
Db	121	isgmchyghghinslykpmldgrdfitalpargldhiaenilslsyldakslcaaelvckewy	180
QY	181	RVTSGMLMKLRIERNVRTDSLMGLAEBRGKQYLFKNKRPDGNAPNSFTYRALVPKIT	240
Db	181	rvtsgmlmwkklriernvrtdslwrglaerirgwgylfknpdpdgnapnsfyralfypkii	240
QY	241	QDIETIESNMWCGHRSIORHICRSETSKGVYCLQYDDQKIVSGRGDRNTIKIMDKNTECK	300
Db	241	qdietiesnwrcgghrsiorhircrsetskgyyclqyddqkivsgrgdrntlikimdkntleck	300
QY	301	RILFGHTGSLVLCLOYDERVITITGSSDSTVRWVDVNTGEMLNTLLIHCEAVLHLRFNNGM	360
Db	301	rilfghtgsvlclqydervilitgssdstvrwvdvntgemlntllihceavlhrlfrnngm	360
QY	361	VTCGSDRSIAVWDMASPDITLRRTVIVGHRNAVVPDDEKIVYSASGDDPTIKVMNSTC	420
Db	361	vtcgsdrrsiavwmaspdiplrtrvlyghraavvpdfrdkivyasaqdrtlkvmnstc	420
QY	421	EEVFTLNGHKRGACIACLOYRDLVYVSSGSSDMNTIRLMDIEGCACLRVLEGHHELVNCIRFDN	480
Db	421	efvrtlnghkrgiacicloyrdrlvvyssgsdmntirrlwdiegcacrlrvleghheylvncirfdn	480
QY	481	KRIVSGAVDGKIKYWDLVYALDPAPAGTCLRTLVHSGRVFRLQDEFQIVSSSHDDT	540
Db	481	krivsgaydgkikywdlvyaaldrpapagtlclrtlvhsgrvfrlqdefqivssshdct	540
QY	541	ILIMDFLNDPAAOAEPPSPSRRTTYISR	569
Db	541	ilimdfindpaagaepprsprrtyltsr	569
RESULT	3		
Y96697			
ID	Y96697	standard; Protein; 569 AA.	
AC	XX		
NC	XX	Y96697;	
DT	XX		
DE	XX	26-SEP-2000 (first entry)	
DE	XX		
DE	XX	Human beta-TRCP.	
XX	XX		
KW	XX	E3 ubiquitin ligase; beta-TRCP; F-box; WD protein; I-kappa-B; inhibitor;	
KW	XX	nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;	
XX	XX	anti-inflammatory; immunosuppressive; cytostatic.	
OS	XX	Homo sapiens.	
PN	XX	WO200034447-A2.	
PN	XX		
PD	XX	15-JUN-2000.	
XX	XX		
PE	XX	10-DEC-1999; 99WO-US29371.	
XX	XX		
PR	XX	10-DEC-1998; 98US-0210060.	
PA	XX	(SIGN-) SIGNAL PHARM INC.	
PA	XX	(YISS) YISSUM RES & DEV CO.	
PI	XX	Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;	
PI	XX	*Lavon I, Yaron A;	
XX	XX		
XX	XX	WPI; 2000-431294/37.	
XX	XX		

Query Match	Best Local Similarity	100.0%	Score 3034;	DB 21;	Length 569;
Matches 569;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY 1	MDPAEVLQEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLINDFTVCLA 60				
DB 1	mdpaevlqekalkfmsnserecdnngpprklipeknsrlrqtynscarlindgtvcla 60				
QY 61	STAMKTEENCVAKTAKLANGTSSMTIVKQKKLASIYEKEELCYKTFEQMSDQVEFVHL 120				
DB 61	stamktencvaktaklangtssmivkqkqliasysykekelcvkyfegsedsdqvfevhl 120				
QY 121	ISOMCHYOHGHINSLTKMLORDFTTALPARLDHIANIITSYDAKSLCAELCKRMW 180				
DB 121	isgmchyqghlnshyikpmlqrdftltparldhlaenllslydaakslcaaelvckewy 180				
QY 181	RVTSDGMLMKLIEEMVVTSDSLMRGLAERRGWGLFKFNKPPDGNAPNSEFYRALYPKII 240				
DB 181	rvtsdgmllmkliemvvtstdslwrglaerwggylfkfkppdgnapnsfyralypkii 240				
QY 241	ODIETIESNMRCGRHSIORHCRSFSGKYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 300				
DB 241	qdietiesnmrcgrhsiorhcrsfsgkyvclqyddqkivsglrdrntikimdkntleack 300				
QY 301	RILHTGTSVLCLOYDERVIITGSSDSIVRWVDVVTGEMLNTLIHHCVAVLHLRPNNGMM 360				
DB 301	riltghtsvlclqydeviitgssdsivrwvdvvtgcmllntlihhcvaavlhlfrnngmm 360				
QY 361	VTCNSDRSLAVWDMASPPDITLRVLVGHRAVNVVDQDKRYVASASDRTIKVWNTSTC 420				
DB 361	vtcnsdrslavwdmasspdiltlrylvghraavnvdqdkryvasasdrtikvwntstc 420				
QY 421	EEFVRLNGHKRIGLQYRDLVWGSSSDNTIRLMDICGACLRVLECHEBELVRCIRSDN 480				
DB 421	eefvrlnghkrglqyrdrlvwgsssdntirlmdicgacrlrvlechebelvrcirsdn 480				
QY 481	KRISGAYDGKTRKVDVLAALDPRAPAGTCLRTLVHSGRYFLQDFEPQIVSSSHDT 540				
DB 481	krivsgaydgktrkvdvllaaldprapagtlclrtlvhsgryflqdfefqvssshdtt 540				
QY 541	ILIMDELNDPAAOAEPPRSPSTTYVIR 569				
DB 541	ililwiflndpaaqaepprspstlycysir 569				

Y83041
 ID Y83041 standard; Protein: 569 AA.
 AC Y83041:
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE F-box protein FBP-1.
 XX
 KW F-box protein: FBP; diagnosis: treatment; screening; agonist;
 KW antagonist; proliferative disorder; differentiative disorder;
 KW breast cancer; prostate cancer; ovarian cancer; cancer;
 KW small cell lung carcinoma; immune disorder; cardiovascular disorder;
 KW inflammatory disorder; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200012679-A1.
 XX
 PD 09-MAR-2000.
 XX
 PF 27-AUG-1999; 99WO-US19560.
 XX
 PR 28-AUG-1998; 98US-0098355.
 PR 03-FEB-1999; 99US-0118568.
 PR 15-MAR-1999; 99US-0124449.
 XX
 PA (UYNV) UNIV NEW YORK STATE.
 XX
 PI Chlaur DS, Pagano M, Latres E;
 XX
 DR WPI: 2000-256635/22.
 DR N-PSDB; Z93350.
 XX
 PT Novel nucleic acid for screening compounds useful for treating
 PT proliferative and differentiative disorders such as cancer and immune
 PT disorders comprises sequences encoding ubiquitin ligases
 XX
 PS Disclosure; Figure 3a; 245pp; English.
 XX
 CC Nucleic acids encoding substrate-targeting subunits of ubiquitin
 CC ligases with F-box motifs (F-box proteins) are useful for diagnosis
 CC of proliferative and differentiated related disorders by measuring
 CC FBP gene expression. Cells expressing such proteins or
 CC their fragments are useful for screening compounds. The compounds
 CC are agonists or antagonists, which are useful for treating a
 CC proliferative or differentiative disorder in a mammal such as
 CC breast, ovarian and prostate cancer and small cell lung carcinoma
 CC and also for opportunistic infections, immune disorders,
 CC cardiovascular diseases and inflammatory disorders. FBP protein,
 CC analogs, derivatives and their subsequences, anti-FBP antibodies
 CC are also useful in diagnosis of the disorders.
 XX
 SO Sequence 569 AA.

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 2.4e-287;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEVLOEKALKFENSSREDCCNNGEPPRKTIPEKNSLRQTYSCARLCLNQTVCIA 60
 DB 1 mdpaeavlgekalikfmsseredcnngepprkilpeksnlrqtyscalclngctvcia 60
 QY 61 STAMKTEVCVARTKLANGSSMIVPQKRLASYEKEKLCVKEQWSESQVFEVHL 120
 DB 61 stamktevcvartklangssmivpqrklissayeke1cvkyeqwesqvefvehl 120
 QY 121 ISOMCHYQGHINSYLPMLQDFITALPARGLDHAENILSYLDAKSLCAAEVCKEMX 180
 DB 121 isgmchycghinsylpmlqdfitalparglidhaenilisyldaksicaaelvckey 180
 QY 181 RVTSGMLMKKLIERNVRTDSLWRGLAERRGGOYLFFKNKPPDGNA PPSFYRALYPKII 240

DB 181 rvtsgmlwkkliernvrtdslwrglaeerrwgqylfknkppdgna ppsfyralypkii 240
 QY 241 ODITFESNMRCGRHSIORHCRSEPSKGYVCLQYDOKIVSGLRONTJIKMDKNLECK 300
 DB 241 oditfesnrcgrhsiorhcrsepskgyvclyyddqkivsglrvnclikmdknleek 300
 QY 301 RILGHTGSVLCLOYDERVITGSSDSTVVRWMDVNTGEMLNTLIHCEAVLHLRFNNGMA 360
 DB 301 rilghgsvlclqyderviltgssdstvrvwmdvntgemlntlihceavhlrfrnngm 360
 QY 361 VTGSKDMSIAVMDASPTDITLRRLVGHRAAVNVDFDDKYIVASGDRTIKWNFTSTC 420
 DB 361 vtcskdrsiavmdasptdiltlrtrlvghraavnvdfddkyivasgdrtilkwnftstc 420
 QY 421 EFVRTLNGHRTACLOYRRRLVSSSDNTIRLMEIEGACLRVLEGHELVRCIRPN 480
 DB 421 efvrtlnghrtacloyrrrlvsssdntirlmiegacrlrvleghelevrcirfn 480
 QY 481 KRIISGAVDGKIKWMDLVAALDPRAPAGTLCRLTIVEHSGRVERLOPDEFQIVSSHDDT 540
 DB 481 krtisgavdgkikwmdlvaaldprapagtlcrltlvehsgvrverlqdefqivsshdtd 540
 QY 541 ILIWDPLNDPAQAEPSPSRRTYTYISR 569
 DB 541 ilwdfplndpaqaepspsrtytyisr 569

RESULT 5
 Y83250
 ID Y83250 standard; Protein: 569 AA.
 AC Y83250:
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE F-box protein hbeta1rcp.
 XX
 KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;
 KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;
 KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO200022110-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US23705.
 PF 09-OCT-1998; 98US-0103787.
 PR
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Zhou P, Howley P;
 XX
 DR WPI: 2000-317970/27.
 DR N-PSDB; Z93710.

Targeting degradation of polypeptide useful for treating cancer and
 other proliferative disorders, involves conjugating polypeptide with
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic
 PT compound
 XX
 PS Claim 9; Page 171; 185pp; English.
 XX
 CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 CC ligases) which can be used for the targeted degradation of a target
 CC polypeptide in vivo. Targeted degradation is achieved by expressing
 CC the ubiquitin ligase in a cell linked to the interaction domain of
 CC the target polypeptide and thereby recruiting the target polypeptide
 CC to the ubiquitin ligase. Such methods are useful for decreasing or

	OS	Homo sapiens.
XX	Key	Location/Qualifiers
XX	Modified-site	19 /note= "Potential phosphorylation site"
FT	Modified-site	39 /note= "Potential phosphorylation site"
FT	Modified-site	91 /note= "Potential phosphorylation site"
FT	Modified-site	109 /note= "Potential phosphorylation site"
FT	Modified-site	162 /note= "Potential phosphorylation site"
FT	Modified-site	266 /note= "Potential phosphorylation site"
FT	Modified-site	288 /note= "Potential phosphorylation site"
FT	Modified-site	328 /note= "Potential phosphorylation site"
FT	Modified-site	376 /note= "Potential phosphorylation site"
FT	Modified-site	381 /note= "Potential phosphorylation site"
FT	Modified-site	411 /note= "Potential phosphorylation site"
FT	Modified-site	418 /note= "Potential phosphorylation site"
FT	Modified-site	431 /note= "Potential phosphorylation site"
FT	Modified-site	514 /note= "Potential phosphorylation site"
FT	Modified-site	519 /note= "Potential phosphorylation site"
FT	Modified-site	535 /note= "Potential phosphorylation site"
FT	Modified-site	536 /note= "Potential phosphorylation site"
FT	Modified-site	17 /note= "Potential phosphorylation site"
FT	Modified-site	77 /note= "Potential glycosylation site"
FT	Modified-site	416 /note= "Potential glycosylation site"
FT	Modified-site	320..334 /note= "Potential glycosylation site"
FT	Region	360..374 /label= Signature_sequence
FT	Region	403..417 /label= Signature_sequence
FT	Region	443..457 /label= Signature_sequence
FT	Region	483..497 /label= Signature_sequence
FT	Region	532..546 /label= Signature_sequence
XX	WO958558-A2.	
XX	18-NOV-1999.	
XX	13-MAY-1999;	99WO-US10567.
XX	13-MAY-1998;	98US-0085343.
XX	26-AUG-1998;	98US-0098010.
PA	(INCY-) INCYTE PHARM INC.	
PI	Bandman O. Hillman JL, Lal P, Yue H, Tang YT, Patterson C,	
PI	Baughn MR, Yang J;	
XX	WPI; 2000-086432/07.	
RR	N-PSDB; 229233.	

```

XX Human cell signaling proteins useful for, e.g. diagnosing cell
PT proliferative and inflammatory disorders
XX
XX Claim 1; Page 77-78; 90pp; English.
XX
XX The present sequence is cell signalling protein-12 (CSIGP-12) encoded
CC by cDNA obtained from Incyte clone 3339149 of COLAUC701 library. It is
CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is
CC found to be homologous to beta-transducin repeats containing
CC protein. Fragments of CSIGP encoding nucleic acid can be used as
CC hybridisation probe for detecting CSIGP related sequences or allelic
CC variants. Recombinant CSIGP can be produced in host cells by transforming
CC them with genetically engineered vectors. Agonists or antagonists can be
CC used in the treatment of cell proliferative and inflammatory disorders
CC associated with decreased or increased CSIGP expression. CSIGP is used in
CC the diagnosis, prevention and treatment of cell proliferative disorders
CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.
XX
XX Sequence 569 AA:
SQ

```

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 2.4e-287;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MDPAEAVLQEKALKFPMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60
DB 1 mdpaeavlgkalkfpmnserecdnngpeprkiipeknslrtyngscarlclngetvcla 60
OY 61 STAKTENCVAKTKLANGTSSMIVPKORKLASYEKEKELCVKFEQWSESDOVEVEHL 120
DB 61 staktencvaktklangtssmivpkorklasyekekeltcvkfeqwesdqvfevhl 120
OY 121 ISQMCYOHGHINSYKPKMLQRFDTALPARGLDHAENILSYLDAKSLCAAEVCKEY 180
DB 121 isqmcyqhghinsylkpmldrfdtalparglidhaenilsyldakslcaaelvckey 180
OY 181 RVTSDGMLMKKLIERNVRFDSLRGLAERRGQOYLFFKNKPPDGNAPNPSFRALYPKII 240
DB 181 rvtsgdmlkkliernvrfdsllrglaerrgwgyilfknpdpnappnsfyralkypkii 240
OY 241 QDIETTESNMRCGRHSIORHCRSETSKGYCLOYDDOKIVSGLRDNTIKIMDKNTLECK 300
DB 241 qdiettesnmrcgrhsiorhcrsetskgyclgyddqkivsgldntikimdkntleck 300
OY 301 RILTGHTGSVLCLOYDERVITIGSSDSYVRVDVNTGEMLNLIIHCEAVLHLPFNNGMM 360
DB 301 rilghtgsvlclgydervilitssdsstvrwvntgemlnliihceavlhlpfnngmm 360
OY 361 VTCSKRSIAVWDMASPTDITLRRLVGHRAAVNVDPDDKIVVASGRTIKVWNTSTC 420
DB 361 vtcskrsiavwmdasptditlrrvlghraavnvdpddkivvasgrrtikvwnstsc 420
OY 421 EFVRLTNGHRRGIACLOYDRVLVSSGSSDNTIRLMDIEGACLRVLEGHLELVRICRFDN 480
DB 421 efvrltnghrgrgiacloydrvlvssgssdntirlmdiegcacrlrvleghlelvrircrfdn 480
OY 481 KRIVSAGYGGKIKVMDLVLAALDPAPAGTLCRTIVHSGRFRLOPDEQIVSSSHDT 540
DB 481 krivsagygkikvmdlvlaaldpapagtlcrtivhsgvrfirldfdefqvssshdt 540
OY 541 IILWDFLNDPAQAEPSPSRITYISR 569
DB 541 iilwdflnmpaqaepspsrityistr 569

```

RESULT 7
 B12812
 ID B12812 standard; protein: 569 AA.
 XX
 AC B12812;

```

XX 27-NOV-2000 (first entry)
DT
XX Mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.
DE
XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;
KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;
KW gene therapy; colon cancer; beta-transducin repeat containing protein;
KW beta-Trcp.
XX
OS Mus musculus.
PN JP2000166542-A.
XX
PD 20-JUN-2000.
XX
PF 02-DEC-1998; 98JP-0343437.
XX
PR 02-DEC-1998; 98JP-0343437.
XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR WPI: 2000-485550/43.
DR N-PSDB; A73131.
XX
PT F-box protein of ubiquitin ligase SCF complex which promotes the
PT ubiquitination of Ikappab or beta-catenin
XX
PS Claim 2; Page 9-10; 19pp; Japanese.
XX
XX The present invention describes an F-box motif protein of ubiquitin
CC ligase SCF complex which promotes the ubiquitination of Ikappab or
CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
CC complex (SCF complex) of F-box protein containing F-box motif and WD40
CC repeat motif and has the amino acid sequence of 45 residues (B12811)
CC or one of two 569 residue sequences (B12812, which is mouse ubiquitin
CC ligase FWD1 protein) and (B12813, which is human beta-transducin repeat
CC containing protein (beta-Trcp)). The F-box protein can be used for the
CC gene therapy of colon cancer by being recombined to a virus vector.
XX
XX Sequence 569 AA:
SQ

```

Query Match 98.8%; Score 2997; DB 21; Length 569;
 Best Local Similarity 98.6%; Pred. No. 9.8e-284;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

OY 1 MDPAEAVLQEKALKFPMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60
DB 1 mdpaeavlgkalkfpmnserecdnngpeprkiipeknslrtyngscarlclngetvcla 60
OY 61 STAKTENCVAKTKLANGTSSMIVPKORKLASYEKEKELCVKFEQWSESDOVEVEHL 120
DB 61 staktencvaktklangtssmivpkorklasyekekeltcvkfeqwesdqvfevhl 120
OY 121 ISQMCYOHGHINSYKPKMLQRFDTALPARGLDHAENILSYLDAKSLCAAEVCKEY 180
DB 121 isqmcyqhghinsylkpmldrfdtalparglidhaenilsyldakslcaaelvckey 180
OY 181 RVTSDGMLMKKLIERNVRFDSLRGLAERRGQOYLFFKNKPPDGNAPNPSFRALYPKII 240
DB 181 rvtsgdmlkkliernvrfdsllrglaerrgwgyilfknpdpnappnsfyralkypkii 240
OY 241 QDIETTESNMRCGRHSIORHCRSETSKGYCLOYDDOKIVSGLRDNTIKIMDKNTLECK 300
DB 241 qdiettesnmrcgrhsiorhcrsetskgyclgyddqkivsgldntikimdkntleck 300
OY 301 RILTGHTGSVLCLOYDERVITIGSSDSYVRVDVNTGEMLNLIIHCEAVLHLPFNNGMM 360
DB 301 rilghtgsvlclgydervilitssdsstvrwvntgemlnliihceavlhlpfnngmm 360
OY 361 VTCSKRSIAVWDMASPTDITLRRLVGHRAAVNVDPDDKIVVASGRTIKVWNTSTC 420
DB 361 vtcskrsiavwmdasptditlrrvlghraavnvdpddkivvasgrrtikvwnstsc 420

```


Db 361 vtcskdrsiawmdasptdlrlrvlvghraavnvdfddkyivsagdrlikwntstc 420
QY 421 EFWRLNGHKRGACIQYRDLVYSGSSDNTIRLMDIECGACLRVLEHETLVRCIRPDN 480
|||||
Db 421 efvrlngkhrgiaciqyrdrlvsgssdntirlmdiecgacrlvleghetlvrciridn 480
QY 481 KRIYSGAVDGKIKYWDVLAALDPAPAGTLCRLTVLHSGHVFRLQDFEQIVSSSHDDT 540
|||||
Db 481 krivsgaydgkikwldmaaldprapagtlclrltlvehsgvfrlqdfefqivssshdtt 540
QY 541 ILIWDPLNDPAQAEPSPSRRTYTIYSR 569
|||||
Db 541 ilwdfindpaahaepprsrtyltylsr 569
RESULT 8
Y83254
ID Y83254 standard; Protein; 569 AA.
XX AC Y83254;
XX DT 16-AUG-2000 (first entry)
XX DE F-box protein FWDLP.
XX KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;
KM destabilisation; proteolysis; drug discovery; gene therapy; cancer;
KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
KM mouse; ss.
XX Mus musculus.
XX NC W0200022110-A2.
XX PN 20-APR-2000.
XX PD 08-OCT-1999; 99WO-US23705.
XX PF 09-OCT-1998; 98US-0103787.
XX PR (HARD) HARVARD COLLEGE.
XX PA
XX PI Zhou P, Howley P;
XX DR WPI; 2000-317970/27.
XX DR N-PSDB; 293714.
XX PT Targeting degradation of polypeptide useful for treating cancer and
PT other proliferative disorders, involves conjugating polypeptide with
PT ubiquitin protein ligase or inhibiting ubiquitination using organic
PT compound
XX PS
XX PS Claim 9; Page 184-185; 185pp; English.
XX CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
CC ligases) which can be used for the targeted degradation of a target
CC polypeptide in vivo. Targeted degradation is achieved by expressing
CC the ubiquitin ligase in a cell linked to the interaction domain of
CC the target polypeptide and thereby recruiting the target polypeptide
CC to the ubiquitin ligase. Such methods are useful for decreasing or
CC increasing the level of a target polypeptide and for creating and
CC expressing a destabilized polypeptide which is subjected to SCF
CC mediated proteolysis. Degrading any desired protein in a cell is
CC useful for preventing or treating diseases caused by the presence of
CC abnormal amount of the specific polypeptides, for drug discovery and
CC for gene therapy. Diseases treated include cancer, by degradation of
CC oncoproteins; huntington's disease, other proliferative disorders and
CC microbial infections. The method provides a quick and easy
CC alternative to gene knockout technology. The target polypeptide can
CC be degraded at all stages, or a specific stage, of development in the
CC mature animal.
XX
XX Sequence 569 AA;

Query Match 98.8%; Score 2997; DB 21; Length 569;
Best Local Similarity 98.6%; Pred. No. 9.8e-284;
Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 MDPAAVLQERALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNDETVCOLA 60
|||
Db 1 mdpaaavlqeralkfmnserecdnngpeprkiipeknsrlqtynscaclclngetvcolt 60
QY 61 STAKTENCVAKTKLANGTSSMIYPKOKRLSASTEKELCVKTFEOMSESDYVEFVHL 120
|||||
Db 61 staktencvaktklangtssmiypkqrlksasyekekelcvkfefwseesdyvefvehl 120
QY 121 ISOMCHVOHGHIINSYTLKPMLODFETLALPARGLHIEHNTISYLDASISCAELVCKEMY 180
|||||
Db 121 isgmchyghghinsylkpmldofetlalparglhieenlisyldasicaaelvckemy 180
QY 181 RVTSQGLMKKLIBRMVWRTDSLMBGLAERRGCGYLFRKNKPPDGNAPPNSFYRALYPRKI 240
|||||
Db 181 rvtsqglmkklibrmvwrtdslmbglerrgwgylfrknkppdgnappnsfyralypkii 240
QY 241 QDFTIESNMRCGRHSIQRIHCRSETSKGVYCLQYDQKIVSGIRDNTIKTMDKNTLECK 300
|||||
Db 241 qdftiesnmrcgrhsilqrhcrsetskgyvclqyddqkivsgirldntlikmdkstleek 300
QY 301 RILGHGVSVCLOQYDERVITITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLEFNNGMM 360
|||||
Db 301 rilghgvsvcloqydervittgssdstvrvmvntgemlntliihceavhlhlfnnngmm 360
QY 361 VTCSKDRSIAVWMDASPTDITLRRVLVGHRAAVNVDFDDKYIVSASGDRTIKWNTSTC 420
|||||
Db 361 vtcskdrsiawmdasptdlrlrvlvghraavnvdfddkyivsagdrlikwntstc 420
QY 421 EFWRLNGHKRGACIQYRDLVYSGSSDNTIRLMDIECGACLRVLEHETLVRCIRPDN 480
|||||
Db 421 efvrlngkhrgiaciqyrdrlvsgssdntirlmdiecgacrlvleghetlvrciridn 480
QY 481 KRIYSGAVDGKIKYWDVLAALDPAPAGTLCRLTVLHSGHVFRLQDFEQIVSSSHDDT 540
|||||
Db 481 krivsgaydgkikwldmaaldprapagtlclrltlvehsgvfrlqdfefqivssshdtt 540
QY 541 ILIWDPLNDPAQAEPSPSRRTYTIYSR 569
|||||
Db 541 ilwdfindpaahaepprsrtyltylsr 569
RESULT 9
R85852
ID R85852 standard; peptide; 517 AA.
XX AC R85852;
XX DT 13-SEP-1996 (first entry)
XX DE WD-40 domain-contg. beta-TRCP protein.
XX KW WD40 repeat region; beta-transducin; protein-protein interaction; drug;
KW intracellular signalling; protein kinase C; homology; motif; modulator;
KW receptors of activated protein kinase; enzyme activity; isozyme; human.
XX OS Synthetic.
XX PN W09521252-A2.
XX PD 10-AUG-1995.
XX PF 31-JAN-1995; 95WO-US01210.
XX PR 01-FEB-1994; 94US-0190802.
XX PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Mochly-Rosen D, Ron D;
 XX WPI: 1995-283772/37.
 DR
 XX
 PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the
 PT activity of a protein, eg. protein kinase C, which interacts with a
 PT protein contg. a WD-40 region.
 XX
 PS
 XX Example 5; Page 80-82; 351pp; English.
 CC Proteins R85851-92 are protein which contain at least one WD-40 (also
 CC called beta-transducin homologous) amino acid repeat motifs. The WD-40
 CC regions are involved in protein-protein interactions between proteins
 CC involved in intracellular signaling. An example of such an interaction
 CC is between protein kinase C and receptors of activated protein kinase
 CC (RACK), esp. RACK-1 (R85850). Proteins R85851-82 were isolated based on
 CC homology with beta-transducin, whereas proteins R85882-92 were isolated
 CC based on homology with the WD-40 consensus sequence (R85893). The
 CC proteins were used to construct the peptides R84928-R85063 and
 CC R85786-R85842. The peptides can be used to identify target proteins
 CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of
 CC proteins involved in protein-protein interaction and to screen for drugs
 CC that will affect protein-protein interaction involving WD-40 domains.
 CC
 XX
 SQ Sequence 517 AA;

Query Match 85.1%; Score 2582.5; DB 16; Length 517;
 Best Local Similarity 91.4%; Pred. No. 2,3e-243; Mismatches 8; Indels 31; Gaps 2;

Matches 467; Conservative 7; Mismatches 8; Indels 31; Gaps 2;
 QY 18 SSREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASPMKTEMCVAKTKLAN 77
 DB 13 aserecdnngpprkilteknrlq-----tklan 42
 QY 78 GISSMIVPKRKISASYEKELCVKFEQWSESDQVEFEHLISOMCHYGHTNSYLK 137
 DB 43 gtsmivpqrkisanyekelcvkyfegwsecdqvefehlismchygghntylk 102
 QY 138 PMQORFIALPRGLDHAENILSYLDAKSLCAELVCKEWRPMYSDGMLKKILERNV 197
 DB 103 pmqorftalprglidhaenilsyldakslcsaelvckewrvmysdgmllkklletrnv 162
 QY 198 RTDSLWGLAERKMGQVYFKNPPGNAFPNSFYRALPKTIODIETTESNMGRGHS 257
 DB 163 rtdslwrglaerkgvqylfknppdgktrpnsfyralpkliqdtetlesnmrghs 222
 QY 258 QRHCHSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTSVLCIQYDE 317
 DB 223 qrhcrsetskgyyclqyddqkivsglrdntlikwckntleckrvmgtsvvlclqyde 282
 QY 318 RVIITGSSDSTVWMDVNTGEMNLIIHCEAVLHFRFNNGMVMVCSKDRSTAVWDMASP 377
 DB 283 rviitg-sdstvrvwdvntgemnlilhnceavlhlfrfngmmvltcskdrstlavwdmsa 341
 QY 378 TDTLRLVAVGHRAAVNVDFDKYIVSASGDRITKVMNTSTCEFRYLNGHRRGIAC 437
 DB 342 tdtlrlvlgvhaavnvdfdkyivsasgdrilkwmtstcefrtlnghkrgiac 401
 QY 438 YRRLVYSSGSDNTIRLMDIEGACLRVLEGBELVRCIRFNNKRIIVSAGYGRKRWDL 497
 DB 402 yrlrlvssgsdntirlmdiegacrlvleghbelvrcirfnnkriivsgaygkrlkvw 461
 QY 498 VAALDRAPAGTCLRTLVHSGRVRLQDFEFOIVSSSHDPTILIMDLNRP 550
 DB 462 vaaldrapagtcrltlvhsgrvrlqdfefoivssshdptilimdlndp 514

RESULT 10
 ID Y96696
 XX Y96696 standard; Protein: 542 AA.
 AC Y96696;

XX 26-SEP-2000 (first entry)
 DT
 XX
 DE Human E3 ubiquitin ligase.
 KW E3 ubiquitin ligase; beta-TrCP, F-box; WD protein; I-kappa-B; inhibitor;
 KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
 KW anti-inflammatory; immunosuppressive; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN W020034447-A2.
 XX
 PD 15-JUN-2000.
 XX
 PF 10-DEC-1999; 99WO-US29371.
 XX
 PR 10-DEC-1998; 98US-0210060.
 XX
 PA (SIGN-) SIGNAL PHARM INC.
 PA (YISS) YISSUM RES & DEV CO.
 PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;
 PI Lavon I, Yaron A;
 DR
 DR N-PSDB; A51228.
 XX
 PT Polypeptide enhancing phosphorylated I-kappaB ubiquitination useful for
 PT treating disorder associated with NF-kappaB activation e.g. cancer,
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its
 PT variant
 XX
 PS Claim 1; Page 70-72; 77pp; English.

CC This is human E3 ubiquitin ligase (E3), which is homologous to human
 CC beta-TrCP, an F-box/WD protein family member. E3 enhances ubiquitination
 CC of phosphorylated I-kappa-B, an inhibitor protein of nuclear factor
 CC kappa-B (NF-kappa-B). Understanding I-kappa-B degradation via the
 CC ubiquitin pathway is useful for identifying modulators of this process
 CC for use in treating diseases associated with activation of NF-kappa-B. In
 CC vitro analysis suggests that deletion of the F-box results in a protein
 CC that functions as a dominant negative molecule in vivo. Transient
 CC over-expression of delta-beta-TrCP (a deletion mutant) inhibited the
 CC degradation of endogenous I-kappa-B-alpha in stimulated Jurkat cells,
 CC resulting in accumulation of phosphorylated I-kappa-B-alpha. E3 can be
 CC used to screen for modulators of NF-kappa-B activity. E3 and beta-TrCP
 CC can be used to modulate NF-kappa-B to treat inflammatory diseases,
 CC autoimmune diseases, cancer and viral infections.
 CC
 XX

Sequence 542 AA;

Query Match 78.6%; Score 2384.5; DB 21; Length 542;
 Best Local Similarity 79.0%; Pred. No. 5,8e-224; Mismatches 45; Conservative 48; Mismatches 41; Indels 31; Gaps 7;

QY 1 MDPPEAVIOEKALKFENSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLA 60
 DB 1 mep-dsviedtclaelms-----vp-----tsilgcanlv---esmcsl 36
 QY 61 S-TAMKTEMCVAKTKLANGSTSSMIVPKRKISASYEKELCVKFEQWSESDQVEFEV 118
 DB 37 sclqsmgsyvc---qlngstssvlyvkrpsegnqykexldklyfdqgsedqyefre 93
 QY 119 HLISOMCHYGHTNSYLKFMPLORDFTALPARGLDHIAENILSYLDAKSLCAELVCKE 178
 DB 94 hlismchygghnhsylkpmldgrdftalprglidhaenilsyldakslcsaelvcke 153
 QY 179 MYRVTSDGMLAKKRIERNVRTDSLNRGLAERGMGOVLFKNKPPDGNAPNSFYRALYRK 238
 DB 154 wgrvtsgmlwklleerwrtdslnrglaergmgoylvfknkppdgngaywqylfknrptdg--ppnsfyrylpyk 211

QY 239 I I D I E T I E S N M R C G R H S L Q R I H C R S E T S K G V C L O Y D D Q K I V S G L R D N T I K I W D K N T L E 298
D 212 I I G D I E T I E S N M R C G R H S L Q R I H C R S E T S K G V C L O Y D D Q K I V S G L R D N T I K I W D K N T L E 271
QY 299 C K R L T H T G S V L C L O Y D E R V I T T S S D S T V R W D V N T G E M L N T L I H C E A V L H L R E N G 358
D 272 C L K V L G H T S V L C L O Y D E R V I T T S S D S T V R W D V N T G E M L N T L I H C E A V L H L R E N G 331
QY 359 M A V T C S K D R S I A V W D M A S P D I T L R R V L G H R A A V N V D D K I V S A S G D R T K I W N T S 418
D 332 L m t c a k d r s i a v w d m a s e d i t l r r l v g h r a a v n v d d k i v s a s g d r t k i w n t s 391
QY 419 T C E P E R T L N G H K R G I A C L O Y R D R L V S G S S D N T I R L M D I E C G A C L R V L E G H E L V R C I R F 478
D 392 t c e f v r t l n g h k r g i a c l o y r d r l v s g s s d n t i r l m d i e c g a c l r v l e g h e e l v r c i r f 451
QY 479 D N R I V S G A V D G R I K W D L V A A L D P R A P A G T L C R T L V E H S G R V E R L O P E F O I V S S S H D 538
D 452 d n k r i v s g a v d g r i k w d l v a a l d p r a p a s t l c r t l v e h s g r v e r l o p e f o i v s s s h d 511
QY 539 D T I L I M D F L N D P A A Q A E P R S P S K T Y T Y I S R 569
D 512 d t i l i m d f l n d p a a q a e p r s p s k t y t y i s r 542

RESULT 11

Y22469
ID Y22469 standard; Protein; 626 AA.
AC Y22469;
DT 29-SEP-1999 (first entry)
DE Human gmyc-N-sel-10 protein sequence.
XX
XX Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KM therapy; mammary gland.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX MO9932623-A1.
XX
XX PD 01-JUL-1999.
XX
XX PF 17-DEC-1998; 98WO-US26820.
XX
XX PR 19-DEC-1997; 97US-0068243.
XX
XX PA (PHAA) PHARMACIA & UPJOHN CO.
XX
XX PI Gurney ME, Li J, Pauley AM;
XX
XX DR MPI: 1999-458026/38.
XX
XX DR N-PSDB; X99712.
XX
XX PS New isolated human sel-10 polypeptides
XX
XX PS Example 3; Page 74-77; 91pp; English.
XX
XX CC This sequence is an epitope-tagged version of a human sel-10 protein
XX
XX CC of the invention. The human sel-10 proteins of the invention are isolated
XX
XX CC from hippocampus and mammary gland. The polypeptides can be used to
XX
XX CC alter presenilin function. Compounds which inhibit either the expression
XX
XX CC or the activity of the human sel-10 polypeptides may reverse the effects
XX
XX CC of mutations to presenilin-1 (PS-1) or PS-2, and therefore may be useful
XX
XX CC for the prevention or treatment of Alzheimer's disease.
XX
XX SQ Sequence 626 AA;

Query Match 21.2%; Score 644.5; DB 20; Length 626;
Best Local Similarity 30.1%; Pred. No. 6e-54;

Matches 169; Conservative 103; Mismatches 222; Indels 67; Gaps 15;
QY 8 L O E K A L F M N S S E R E D C N N E P P R K I I P E K - - N S I A R O - - T Y N S C A R L C L N O E T V C I A S T A 63
D 53 m e q k l l s e e d n e m e s y g d l t m e g k l l s e e d n m k k l d h g s e v r f s y g k k p c k s e y 112
QY 64 M K T E N C V - - - - - A K T I A N G T S M I V P K O R K I S A S Y E K E L - - C V K Y E Q M S E S D 112
D 113 t s t g l p c s a t p t f f g d l r a a n g q - - - - - q r r i l t v q p r l g e w l k m f g s e g p e 167
QY 113 Q V E R V E H L I Q M K Y C H G H T N S Y L K P M L O R D F T L A R G I D H A E M I L S Y L D A K S I C A A 172
D 168 k l l a l d e l i s c e p t y g k m m q v l e p q f g d f i s l i p - - - k e l a l y l s f l e p k d l l q a 223
QY 173 E L V C K E W Y R V T S D M L M K K I L E R M V R T D S L M R G L A E - - - - - R R G M G Y L F R K N K P P D G N A P 227
D 224 a q c r y r l l a e d n l l v r e k e - - - - - e g i d e p l h k r r k - - - - - v l p g f i h s p 265
QY 228 P N S F Y R A L Y F K I I O D I E T I E S N M R C G R H S L Q R I H C R S E T S K G V C L O Y D D Q K I V S G L R D N 287
D 270 w k s a y - - - - - i r q - - h r i d l n v r t g e l k s p k v - l k y h d h v l t c l q f c g n r l v s g s d d n 320
QY 288 T I R I W D K N T L E C K R I L N G H T S V L C L O Y D E R V I T T S S D S T V R W D V N T G E M L N T L I H C 347
D 321 c l k w s a v t g k c l t l v g h t g y v s g m r d n l l s g d t l k w m a e t g e c i n t l y g h t 380
QY 348 E A V L H L R F N N G M A V T C S K D R S I A V W D M A S P D I T L R R V L G H R A A V N V D D K I V S A S 407
D 381 s t y r c m l h k e k r v s g s a d e l l r w d i e t g c l - - - h v l m g h a a v r c v g d g r i v v s g a 437
QY 408 G D R T I K W N T S T C E F V R T L N G H K R G I A C L O Y R D R L V S G S S D N T I R L M D I E C G A C L R V L E 467
D 438 y d f m v k w d p e t c l h t l g h t n r y s l q f d g l h v v s g l d e s i r w d v e t g n c i n t l t 497
QY 468 G H E L V R C I R F D N K R I V S G A V D G R I K W D L V A A L D P R A P A G T L C R T L V - - - E H S G R V R 524
D 498 g h g s l t s g m e l k a n l y s g n a d s t v k i w d l t g - - - - - g c l q t l g g p n k h q s a v t c 548
QY 525 L Q F D E F Q I V S S H D D T L I M D 545
D 549 l q f d e f q i v s s d d g t v k l w d 569

RESULT 12

B59201
ID B59201 standard; protein; 626 AA.
AC B59201;
DT 23-MAR-2001 (first entry)
DE Protein encoded by construct gmyc-N-sel-10.
XX
XX
XX Sel-10; human; Alzheimer's disease; Abeta.
KM
XX
XX Synthetic.
XX
XX MO200075328-A1.
XX
XX PD 14-DEC-2000.
XX
XX PF 23-MAY-2000; 2000WO-US09814.
XX
XX PR 09-JUN-1999; 99US-0328877.
XX
XX PA (PHAA) PHARMACIA & UPJOHN CO.
XX
XX PI Pauley AM, Gurney ME, Li J;
XX
XX DR MPI: 2001-102404/11.
XX
XX PT New human sel-10 polypeptides and their encoding polynucleotides,
XX
XX useful for raising antibodies for detecting sel-10 polypeptide

Db 431 nadvsvkikwdikt-----gcqlqlgpnkqhsavtclqfknfvitssdgtvkl 481
Qy 544 WD 545
Db 482 wd 483

RESULT 14
Y22468
ID Y22468 standard; Protein; 540 AA.
XX
AC Y22468;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human mammary sel-10 protein sequence.
XX
KW Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;
KW mammary gland; therapy.
XX
OS Homo sapiens.
XX
PN MO9932623-A1.
XX
PD 01-JUL-1999.
XX
PF 17-DEC-1998; 98WO-US26820.
XX
PR 19-DEC-1997; 97US-0068243.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Gurney ME, Li J, Pauley AM;
XX
DR WPI; 1999-458026/38.
XX
DR N-PSDB; X99702.
XX
PS New isolated human sel-10 polypeptides
XX
PS Claim 24; Page 66-69; 91pp; English.
XX
XX
CC This sequence represents a human sel-10 protein of the invention. This
CC sequence is specifically a human mammary sel-10 protein. The polypeptides
CC can be used to alter presenilin function. Compounds which inhibit either
CC the expression or the activity of the human sel-10 polypeptides may
CC reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and
CC therefore may be useful for the prevention or treatment of Alzheimer's
CC disease.
XX
XX
SQ Sequence 540 AA;

Query Match 21.1%; Score 640; DB 20; Length 540;
Best Local Similarity 30.8%; Pred. No. 1.3e-53;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

Qy 14 KFMNSEREDCNNGEPPKRIIEKNSLRQYNSCARLCLNOFTVCLASTAMKENCVAKT 73
Db 4 klhgservrslgkpkv-----seyrtstgl-----vpsca-----tpftfgdl 45
Qy 74 KLANGTSSNIYVKKOKKLSAYEKEL--CYKYEQMSBSDOVEFVEHLISOMCHYQHGH 131
Db 46 raanqgg-----qgrritlsyqptglqewlkmfqswepekllaldelldscptqvk 100
Qy 132 INSYLKPMLORFITALPARGDHAENILSYLDAKSLCAAEVLCKEWRVRYSDGMLMKK 191
Db 101 mmqvlepfqfqlslp---keialyvlstlepkdlldaaqctrryrlaencllwre 156
Qy 192 LIERVVRTDSLMRGLAE-----RGMGOYLFEKNKPPDGNAPNSFYRLYPIIIDIEFI 246
Db 157 kcke-----egldelphlkrirk-----vfkpgflhspsksay-----lrg-hri 194
Qy 247 ESNMHCGRHSIORICRSETSKGYCLOYDOKIVSGLRDNTIKIMDKNTLECKRILTGH 306

Db 195 dtuwrregelkspkv-lkghddhvtclclqfagnrlyvsaddnllkwsavtgkclrtlvgh 253
Qy 307 TGSVLCLOYDERVITITGSDSTVRWVDVTGEMLNTLIHCEAVLHLRPNMAMTCSGD 366
Db 254 tggvssqmdnllsgstdrllkxwaaetgcehllyghstvcvcmllhkrvvsgrsd 313
Qy 367 RSIAYMDASPTDITRLRYLVGHRAAVNVDPDKYIYASASGDRITKWNSTCEFRVL 426
Db 314 atlrwvdieltgcl--hvlmgvaavrcvgydgrtrvvsaydfmfwkwpdetcclhl 370
Qy 427 NGHKRGIACTQYRDRLVYSGSSDNTIRLMDICGACRLVBESHELVKRCIRDNKRIYSG 486
Db 371 qghlnrvyslqfdgllhvsgrsltslrwvdeitgncihltgqsltsamejkdnllyvs 430
Qy 487 AYDGRIKYWDLYAALDPRAPAGTLCRLTV---EHSGRVFRLOPFQFOIVSSHDDTILI 543
Db 431 nadvsvkikwdikt-----gcqlqlgpnkqhsavtclqfknfvitssdgtvkl 481
Qy 544 WD 545
Db 482 wd 483

RESULT 15
B59197
ID B59197 standard; protein; 540 AA.
XX
AC B59197;
XX
DT 23-MAR-2001 (first entry)
XX
DE Human hippocampal sel-10-5 protein.
XX
KW Sel-10; human; Alzheimer's disease; Abeta.
XX
OS Homo sapiens.
XX
PN WO200075328-A1.
XX
PD 14-DEC-2000.
XX
PF 23-MAY-2000; 2000WO-US09814.
XX
PR 09-JUN-1999; 99US-0328877.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Pauley AM, Gurney ME, Li J;
XX
DR WPI; 2001-102404/11.
XX
PS New human sel-10 polypeptides and their encoding polynucleotides,
PS useful for raising antibodies for detecting sel-10 polypeptide
PS expression and as drug targets in the treatment of Alzheimer's disease
PS
XX
XX
CC Claim 1; Page 75-78; 116pp; English.
XX
CC The present invention relates to human sel-10. The sel-10 proteins of
CC the invention are useful for raising monoclonal or polyclonal
CC antibodies useful in diagnostic assays for detecting sel-10
CC polypeptide expression. The sel-10 polypeptides are also useful as drug
CC targets for decreasing antibody levels in the treatment of Alzheimer's
CC disease. It is also useful for identifying agents capable of
CC altering the production level of Abeta. The polynucleotides are useful
CC for developing assays for identifying agents capable of interfering
CC with the biological pathways that lead to Alzheimer's disease.
XX
XX
SQ Sequence 540 AA;

Query Match 21.1%; Score 640; DB 22; Length 540;

Best Local Similarity 30.8%; Pred. No. 1.3e-53;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

```
OY 14 KPMNSSEEDCNGEPPKIIPEKNSLRQTYNSCARCLNQETVCLASTAMKTENCVAKT 73
Db 4 kldhsevrstslgkpkckv-----seytstgl-----vpcsa-----tpitfgdl 45
OY 74 KLANGTSMTIVPKORLSAYEKREL--CVKYPQWSESDQVEFEVHLISQMCHYOHGH 131
Db 46 raanggg-----qgrritsvppptglgwlkmgfswsgpekllaldeldsceptgvkh 100
OY 132 INSYLKPMLODFITALPARGLDHAENILSYLDAKSLCAELVCKEMYRVTSDEMLWKK 191
Db 101 mmqvlepfqgrdfisllp-----kelalylsflepkdl1qaagctcrywrlaeadhllwre 156
OY 192 LIERAVRTSLSMRGLAE-----RGMGOYLFRKNRPPDGNAPRNSFYRALYPKIIDIERI 246
Db 157 kcke-----egldephllkrk-----vfkpgflhspxksay-----lrq--hrl 194
OY 247 ESNMRCGRHSLORIHCRSETSKGVYCLQYDDQKIYSGLRDNTIKIMDKNTLECKRILTGH 306
Db 195 dcnwrrtgelkspkv--lkghddhviclqfcgnriyvsdndtlkxwsavtgcrltlvgh 253
OY 307 TGSVLCLOYDERVITIGSSDSTFVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMNVTCSDK 366
Db 254 tgvvssqmrndhllsgstcdtlkxwnaetgecithlyghtstvtzcmhlhekrvvsgrd 313
OY 367 RSLAVWDMASPTDITLRRLVGHRAAVNVDPDDKIYASGDRTIKVNNTSTCEFEVRTL 426
Db 314 atlrwddietgcgl---hvlmgvvaavrcvqydgrrvvsgaydfmfvkvwdepetclhll 370
OY 427 NGHKRGIAILOQRDLRVVSSSDNTIRLMDIEGACLRVLEGHHEELVRCIRFDRNRIVSG 486
Db 371 qghtrrvyqlqfdghlvsgsldstlrwddetgncihcltghqsltsgmelkdhllvsg 430
OY 487 AYDGKIKVMDLYAALDPRAPAGTLCRLTV---EHSGRVFRLOQDFEQIVSSSHDDTILI 543
Db 431 nadstvkikwdikty-----gcilqtlgpnkqhsavtclqfnknfvtltsddgtvkl 481
OY 544 WD 545
Db 482 wd 483
```

Search completed: June 7, 2001, 10:35:31
Job time: 30 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 10:35:01 ; Search time 13.8 Seconds
(without alignments)
792.100 Million cell updates/sec

Title: US-09-415-795-4

Perfect score: 3034
Sequence: 1 MDPAEVLDKALKFNNSSE.....PAAQADPPSPRTYYISR 569

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2582.5	85.1	517	1 US-08-190-802A-30	Sequence 30, Appl.
2	520	17.1	587	3 US-08-899-578-2	Sequence 2, Appl.
3	399	13.2	779	1 US-08-190-802A-32	Sequence 32, Appl.
4	354	11.7	409	2 US-08-283-917-3	Sequence 3, Appl.
5	354	11.7	409	2 US-08-961-716-3	Sequence 3, Appl.
6	354	11.7	410	2 US-08-283-917-9	Sequence 9, Appl.
7	354	11.7	410	2 US-08-961-716-9	Sequence 9, Appl.
8	339.5	11.2	409	1 US-08-190-802A-51	Sequence 51, Appl.
9	321.5	10.6	514	1 US-08-190-802A-66	Sequence 66, Appl.
10	318	10.5	422	1 US-08-190-802A-52	Sequence 52, Appl.
11	313.5	10.3	209	3 US-08-899-578-6	Sequence 6, Appl.
12	306	10.1	704	1 US-08-188-582-5	Sequence 5, Appl.
13	306	10.1	704	1 US-08-646-715-5	Sequence 5, Appl.
14	305	10.1	704	1 US-08-190-802A-62	Sequence 62, Appl.
15	305	10.1	704	2 US-08-308-818-3	Sequence 3, Appl.
16	298.5	9.8	704	1 US-08-188-582-18	Sequence 18, Appl.
17	298.5	9.8	704	1 US-08-646-715-18	Sequence 18, Appl.
18	285.5	9.4	798	1 US-08-190-802A-64	Sequence 64, Appl.
19	285.5	9.4	798	1 US-08-190-802A-68	Sequence 68, Appl.
20	285.5	9.4	798	1 US-08-308-818-2	Sequence 2, Appl.
21	283.5	9.3	212	3 US-08-899-578-7	Sequence 7, Appl.
22	283	9.3	713	1 US-08-190-802A-63	Sequence 63, Appl.
23	263	8.7	318	1 US-08-190-802A-33	Sequence 33, Appl.
24	256	8.4	906	1 US-08-190-802A-31	Sequence 31, Appl.
25	245.5	8.1	317	1 US-08-190-802A-27	Sequence 27, Appl.
26	245.5	8.1	317	1 US-08-190-802A-41	Sequence 41, Appl.
27	245.5	8.1	317	1 US-08-190-802A-47	Sequence 47, Appl.

28	239	7.9	2627	2 US-08-751-189-3	Sequence 3, Appl.
29	239	7.9	2627	2 US-09-060-836-3	Sequence 3, Appl.
30	239	7.9	2627	4 US-09-184-445-3	Sequence 38, Appl.
31	236	7.8	340	1 US-08-190-802A-38	Sequence 42, Appl.
32	235	7.7	340	1 US-08-190-802A-42	Sequence 1, Appl.
33	233	7.7	395	3 US-09-032-372-1	Sequence 39, Appl.
34	232.5	7.7	326	1 US-08-190-802A-39	Sequence 4, Appl.
35	226	7.4	2629	2 US-08-751-189-4	Sequence 4, Appl.
36	226	7.4	2629	2 US-09-060-836-4	Sequence 4, Appl.
37	226	7.4	2629	4 US-09-184-445-4	Sequence 4, Appl.
38	222	7.3	305	3 US-08-965-600-1	Sequence 1, Appl.
39	215	7.1	340	1 US-08-190-802A-40	Sequence 40, Appl.
40	208.5	6.9	341	1 US-08-190-802A-45	Sequence 45, Appl.
41	203	6.7	350	2 US-08-828-922-1	Sequence 1, Appl.
42	196.5	6.5	439	1 US-08-190-802A-65	Sequence 65, Appl.
43	194.5	6.4	325	1 US-07-626-589-2	Sequence 2, Appl.
44	194.5	6.4	325	1 US-08-190-802A-57	Sequence 57, Appl.
45	194.5	6.4	325	1 US-08-323-444A-1	Sequence 1, Appl.

ALIGNMENTS

RESULT 1
US-08-190-802A-30
Sequence 30, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190, 802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33, 875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-190-802A-30

Query Match 85.1%; Score 2582.5; DB 1; Length 517;
Best Local Similarity 91.4%; Pred. No. 4, 7e-266;
Matches 487; Conservative 8; Indels 31; Gaps 2;

```

01 18 SSEPDCNNGPPKIIPEKKSILQYNTSARCLQNEFTVCLASTAMKTECNCAKTKLAN 77
02 13 ASERDCRDEPPKIIITENTNTLQ-----TKLAN 42
03 78 GTSSMIVPKOKKSASYEKEKELCVKFEQWMSSEDOVEFVHLISOMCHYOHGHSYUK 137
04 43 GTSSMIVPKOKKSANTEKEKELCVKFEQWMSSEDOVEFVHLISOMCHYOHGHSYUK 102
05 138 PMLORDFTALPANGLDHIAENILSIYDAKSLCAELVCKEWTAVTSDGMLMKKILIERAV 197
06 103 PMLORDFTALPANGLDHIAENILSIYDAKSLCAELVCKEWTAVTSDGMLMKKILIERAV 162
07 198 RTDSIMRGLARBRMGOLFENKRPDPCNAPNSPYRALYKIIODIETTESNRGGRHSI 257
08 163 RTDSIMRGLARBRMGOLFENKRPDPCNAPNSPYRALYKIIODIETTESNRGGRHSI 222
09 258 QRHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSVLCQYDE 317
10 223 QRHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSVLCQYDE 282
11 318 RVIITGSSDSVRYWVDVNTGEMLTLLIHCEAVLHLRFNNGMAYTSKDRSIYVWMDASP 377
12 283 RVIITG-SDSVRYWVDVNTGEMLTLLIHCEAVLHLRFNNGMAYTSKDRSIYVWMDASA 341
13 378 TDTILRRLVGHRAAVNVDDQKYIVASGDRITKVMNTSTCEPFTLNGHGRGIACIQ 437
14 342 TDTILRRLVGHRAAVNVDDQKYIVASGDRITKVMNTSTCEPFTLNGHGRGIACIQ 401
15 438 YRDLRVGSSDNTIRLMDIECGACLVLESGHELVYLCIRFDMKRIYSGAYDGIKIVMDL 497
16 402 YRDLRVGSSDNTIRLMDIECGACLVLESGHELVYLCIRFDMKRIYSGAYDGIKIVMDL 461
17 498 VALDPRAPACTLCIRLVEHSGFVRLQDFEFOIVSSSHDDFTLLWDFLNDP 550
18 462 VALDPRAPACTLCIRLVEHSGFVRLQDFEFOIVSSSHDDFTLLWDFLNDP 514
19
20 RESULT 2
21 US-08-899-578-2
22 ; Sequence 2, Application US/08899578
23 ; Patent No. 6087153
24 ; GENERAL INFORMATION:
25 ; APPLICANT: Greenwald, Iva
26 ; APPLICANT: Hubbard, E. Jane
27 ; TITLE OF INVENTION: SEL-10 AND USES THEREOF
28 ; NUMBER OF SEQUENCES: 7
29 ; CORRESPONDENCE ADDRESS:
30 ; ADDRESSEE: Cooper & Dunham LLP
31 ; STREET: 1185 Avenue of the Americas
32 ; CITY: New York
33 ; STATE: New York
34 ; COUNTRY: U.S.A.
35 ; ZIP: 10036
36
37 COMPUTER READABLE FORM:
38 MEDIUM TYPE: Floppy disk
39 COMPUTER: IBM PC compatible
40 OPERATING SYSTEM: PC-DOS/MS-DOS
41 SOFTWARE: PatentIn Release #1.0, Version #1.30
42 CURRENT APPLICATION DATA:
43 APPLICATION NUMBER: US/08/899,578
44 FILING DATE: 24-JUL-1997
45 CLASSIFICATION: 435
46 ATTORNEY/AGENT INFORMATION:
47 NAME: White, John P.
48 REGISTRATION NUMBER: 28,678
49 REFERENCE/DOCKET NUMBER: 0575/53200/JPM/AKC
50 TELECOMMUNICATION INFORMATION:
51 TELEPHONE: (212) 278-0400
52 TELEFAX: (212) 278-0525
53 INFORMATION FOR SEQ ID NO: 2:
54 SEQUENCE CHARACTERISTICS:
55 LENGTH: 587 amino acids
56 TYPE: amino acid
57
58

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-899-578-2

```

Query Match	17.1%;	Score 520;	DB 3;	Length 587;
Best Local Similarity	28.8%;	Pred. No. 2e-46;		
Matches 150;	Conservative 78;	Mismatches 222;	Indels 70;	Gaps 17;

```

01 72 KTKIANGNSMI-----VPRGR--LSASYEKEL-----CYKYEOMSESQVE 115
02 35 ESSIYSGSSSYNADKLSSSPHLOKHLDLSASPSRNNDLNRPVBNHLLALFKDLSSAFOMD 94
03 116 FVEHLISOMCHYOHGHINSYKLPMLQRFDTALPARGLDHAENILSYLDAKSICAELV 175
04 95 AFTRLQSSNNTNTRQLRAIIEPHQDFPLSCUPV-----ELGKMLHNLGYDLLKAOV 150
05 176 CKEMYRYTSDGMLKKL-IERMV-----RTDSLHNGLAERRGWGOYLFRKKPPDGNAP 227
06 151 SKNMKLISEIDKIKMSLGVEEFKHHDPDTRVTGAMOGTAIAG-----VTIPDHQP 203
07 228 PN-SYRYALYKIIODT-----ETISNMRGRHSIQRIHCSSEYSKYUCL 273
08 204 CDLVNHRFLKQKFGDIFERRADKSRYLADKLEKNNANPIMGSAV-LRGHEDHYTCM 262
09 274 QYDDOKIYSGRLDNTIKIRWDNKTLECKRIILGHTGSVLCLOYDE--RVITIGSSDSVVR 331
10 263 QIHDDVLTGSDDNLTAKWCIDKKEWMTVLGHGHWGTQSIQCGKRIYSGSDRKYV 322
11 332 MDVNTGEMLNTLIHCEAVLHLRFENNQMWTCSKRSIAVMDASPTDITLRLRYLVGHR 391
12 323 WSTYDGSLLHRTLOGHTSVRCMAMAGSILVTGSRDITLRYWDVESGRHLA---TLGHNA 379
13 332 AVNRYVDFPDKIYVASGDRTIKVWNTSTCEVYRLNHHKGIACIQYRDR--LVYSSSD 449
14 380 AVRQOFGFTGVTVSGYDFTVKINNAHTGRCIFRLTGHNNRNVYSLFESERSYVCSILD 439
15 450 NTIRLMDI---ECGACRYLVEHGEELVRCIFRDNKRIYSGAYDGIKRYMDLVVALDPRAP 506
16 440 TSIKRWDTTREGOCVALLQGHISLTISGMOLKRNILYSCNADSHVAVMDI-----H 491
17 492 BGT-CVHMLSGHRSALITSLQMFGRNMVATSSDDGTVALMD 530
18
19 RESULT 3
20 US-08-190-802A-32
21 : Sequence 32, Application US/08190802A
22 : Patent No. 5519003
23 :
24 : GENERAL INFORMATION:
25 :
26 : APPLICANT: Mochly-Rosen, Daria
27 : APPLICANT: Ron, Dorit
28 : TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
29 : TITLE OF INVENTION: thereof
30 : NUMBER OF SEQUENCES: 265
31 :
32 : CORRESPONDENCE ADDRESS:
33 : ADDRESSEE: Delhinger & Associates
34 : STREET: P.O. Box 60850
35 : CITY: Palo Alto
36 : STATE: CA
37 :
38 : COUNTRY: USA
39 : ZIP: 94306-0850
40 :
41 : COMPUTER READABLE FORM:
42 : MEDIUM TYPE: Floppy disk
43 :
44 : COMPUTER: IBM PC compatible
45 : OPERATING SYSTEM: PC-DOS/MS-DOS
46 : SOFTWARE: Patentin Release #1.0, Version #1.25
47 : CURRENT APPLICATION DATA:
48 : APPLICATION NUMBER: US/08/190,802A
49 : FILING DATE: 01-FEB-1994
50 : CLASSIFICATION: 530
51 :
52 : ATTORNEY/AGENT INFORMATION:

```


NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33, 875
 REFERENCE/DOCKET NUMBER: 8600-0139
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 779 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: Protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: CD4 / CDC20 protein, Fig. 15
 US-08-190-802A-32

Query Match 13.2%; Score 399; DB 1; Length 779;
 Best Local Similarity 24.3%; Pred. No. 2,4e-33;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TYCLASTAKTENCYAKTKLANTGSMIYPRKQKLSASTEKK-----ELCVKTFEOM 108
 DB 194 TPLAKTKTKTINN-----NNNIADLIESKDSIISPELSEISAINNNLPHAYFK-- 244
 QY 109 SSSDOVEFEHLISOMCHYOHGINSYLPMLORDEFTALPARGLDHIAENILSYLDAKS 168
 DB 245 -----NLFLRYVANDSELSDLGLINDKRDITSLP---EISLKFNTYLOFED 294
 QY 169 LCAAEVCEKEMR-VTSDMLMKLI--ERAVRTDSLWGLAERNGWGYLFKNKPPDGN 225
 DB 295 IINSIGVSONMKIIRKSTLKKLISENFV-----SPKGF 331
 QY 226 APPNSFYRLYKTIOD-----IET--IESNRCGRSLDRICRSTSKGYCLOTD 276
 DB 332 NSLNKLKLSOKYFKLSQODRLRISFLENFILKNWYNPKFVPORTLRLGHTVITCLOPE 391
 QY 277 DQKIVSGLRDNITIKIMDKNTLECKRILTGHTGSVYLCLOYDE--RVITGSSDSSTVRWDVN 335
 DB 392 DMYVTGTGADDKMIRYDSTINKKFLQLSGHDGVAALKRAHGILVSGSTDTYTRWMDIK 451
 QY 336 TGEMLNTLIHCEAVLHLRFNNGMNVYTSKDRSLAVWNASPTDITLRLVVGHRAAVN 395
 DB 452 KG-----CCT-----HVFEGHNSVTRC 468
 QY 396 VDFED---KYIVSASGRTIKVNT-----STCE-----PRTLNG 428
 DB 469 LDIVYKNIKITYVTGSRDNTLHWKLPRESSVPDGEHDPYLVFHTPEENPYFVGVLNG 528
 QY 429 HRRGIACLOQRDLVYSGSSDNTIRLMDIEGACLRVLEGEHELVRCLRFDN--KRIYSG 486
 DB 529 HMAVRYSVGHNIYVSSGYDWTILWYDAQMKCLYILSGHTRDRIYSTIYDHEKRCISA 588
 QY 487 AYDGKIKYWDL-----VAALDPRAPAGTL--CLRTLVHSGRVRFLQDFEQIVSS 535
 DB 589 SMDTIRIDNLEINWNGECSYATNSASPCAKITLGAMTLQGTALVGLRLSDREFLVA 648
 QY 536 SHDDTILIMDFLNDPAAQAPRPSRRTTY 566
 DB 649 AADGSIRGMD-AND-----YSRKFYS 668

RESULT 4
 US-08-283-917-3
 ; Sequence 3, Application US/08283917
 ; Patent No. 5849557
 ; GENERAL INFORMATION:
 ; APPLICANT: ADACHI, HIDEKI
 ; APPLICANT: TSUJIMOTO, MASAFUMI
 ; APPLICANT: INOUE, KEIZO
 ; APPLICANT: ARAI, HIROYUKI

TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
 TITLE OF INVENTION: AND GENE THEREOF
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
 ADDRESS: NEUSTADT, P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/283,917
 FILING DATE: 03-AUG-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 209943/1993
 FILING DATE: 03-AUG-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Oblon, No. 5849557/man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2292-030-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 409 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Bos taurus
 US-08-283-917-3

Query Match 11.7%; Score 354; DB 2; Length 409;
 Best Local Similarity 29.2%; Pred. No. 5e-29;
 Matches 87; Conservative 57; Mismatches 110; Indels 44; Gaps 8;

QY 280 IYSGLRDNITIKIMDKNTLECKRILTGHTGSVYLCLOYDE--RVITGSSDSSTVRWDVNTG 337
 DB 122 MVSASEDAITIKWYDETGTGDFERTLGHGTHDSYODISFDHSGKLASCADMTIKLWDFOGF 181
 QY 338 EMLNTLIHCEAVLHLRF--NNGMNVYTSKDRSLAVWNASPTDITLRLVVGHRAAVN 395
 DB 182 ECFIRMHGHDHNVSSVAILMPNGDHIVASRDRTIKMVEYQYTCV---KTFGHEKEMVRM 238
 QY 396 V--DEDDKYIVSASGRTIKVNTSTCEFTLNGHRRGIACLOQRDR----- 441
 DB 239 VAPNDGTLIASCNDQYRVWVAVATKECKALREHENVECISNAPRESSYSISEATGS 298
 QY 442 -----LVYSSSDNTIRLMDIEGACLRVLEGEHELVRCLRFDN--KRIYSGAYD 489
 DB 299 ETKSGKPGPFLISGSRDKTIKMPDVSTGMCMTLVGHDMWVRGVLFSGGKFFILSCADD 358
 QY 490 GKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRFLQDFE--QIVSSSHDDTILIMD 545
 DB 359 KTLRWWDY-----KNKRCMKTLNAHEHVTSLDPRKTAPLYVVTGSVQOTVAYWE 407

RESULT 5
 US-08-961-716-3
 ; Sequence 3, Application US/08961716
 ; Patent No. 5880272
 ; GENERAL INFORMATION:

Db 359 KTLRWMDY-----NKRCKMTLNAHEHFTSLDFHKTAPYVVTGSVDQTVKWE 407

RESULT 9

US-08-190-802A-66
Sequence 66, Application US/08190802A
Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190,802A

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fadian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0960

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:

LENGTH: 514 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49

US-08-190-802A-66

Query Match 10.6%; Score 321.5; DB 1; Length 514;

Best Local Similarity 30.6%; Pred. No. 2.1e-25;

Matches 88; Conservative 44; Mismatches 97; Indels 59; Gaps 9;

Db 236 SIQRHCHSESRKGYCYQYDQKIVSGLRDNTIKIMKNTIECKRIITGHTGSLVLCQY 315

Db 236 SWEPIHLVKKPK-----PRIASSKDGITIKIMPTVSRVCOYTSGHGHSVCVKW 286

Db 316 -DERYITGSSSTYRVMDVNT-GEMLNTLIHHCDAVLHLRPN----- 356

Db 287 GGQGLLYSGSHDRYRVMDINSQRCINILKSHAWVNHLSISTYALRIGAFDHTGKKP 346

Db 357 -----NG---MMVYCSKDRSIAVWMDASPTDITLRRVIVGHRAA 392

Db 347 STPEBAQKALENYEKICKKNGNSEMMVFTASDYTMFLMPLKSTKPLAR--MTGHQKL 404

Db 393 VNVVDF--DDKYIVASASDRITIKYVNTSTCEVYRVLNKHKGICLQYRD--RLVYSSSS 448

Db 405 VNHAVESPDGRITIVASAPDNSIKIMDGDGKFIPTFRGHIAVYQVAVWASSDCRLIVSCSK 464

Db 449 DNTIRLMDIEGACLRVLEG--HEELVRCIRPDNKRIVSGAVDGKIKVW 495

Db 465 DTTLKVMYVTRKRLSVLDLPGIKTKLIYVMSVDGKRVCSGGGDKMVRWL 512

RESULT 10

US-08-190-802A-52
Sequence 52, Application US/08190802A
Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190,802A

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fadian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0960

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 52:

SEQUENCE CHARACTERISTICS:

LENGTH: 422 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: MD6, Fig. 35

US-08-190-802A-52

Query Match 10.5%; Score 318; DB 1; Length 422;

Best Local Similarity 24.4%; Pred. No. 3.0e-25;

Matches 113; Conservative 69; Mismatches 161; Indels 120; Gaps 16;

Db 103 KYFQWSESDVEF-----VEHLISOMCHYOHGHINSYKPMLOPDTALPAR 151

Db 4 KDFETWLDNIVYTELISLMDLQNETLHLISLGAVALRHLSNNLETLTKRDFLKLPL- 62

Db 152 GLDHAENILSYLDAKSICAEALVCKEYRVTSQGLMKLIERMVRVTSIMGLAERRG 211

Db 63 ---ELSFYLLKLPQTLTLCCLVSKQRNVIY-----ACTEWAQFACNMG 106

Db 212 WGVLFKMKPPDGNAPRPSFYRALYPRKIDIDETIESNMGRHSLSGRHRSSTSGVY 271

Db 107 W-----QIDDSVQDSLHWKKVYLKAILRMKQLED-----HEAFETSS--- 143

Db 272 CLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSLVLCQYDERVITGSSDSTVRV 331

Db 144 -----LHISARVVALYKKDGLCTGSDLSAKL 172

Db 332 WDVNTGEMLNTL-IHHCDAVLHLRPNNGMAYTCSKDRSIAVWMDASPTDITLRRVIVGHR 390

Db 173 WDVSTGCVVYGIQHTCAAV---KFDEQKIVTGSFDMTVACMEWSSGARQDHFH---GHT 226

Db 391 AAVNVVDFDDK--YIVASGDRITIKVYVNTSTCEVYRVLNKHKGRI-----ACLO 437

Db 227 GAVESVYDEDELDILVSSGADPAVKVWALSGICLNTLGHTEWTKVYLQCKVSLH 286
QY 438 YRDLVYSSSDNTIRLW---DIEGACLRVLEGHFELVRCIR---FDNKRIYS---- 485
Db 287 SPEDYILLSADKREIKIMPIGREINC-KCLKTLVSSEDSICLOPRLHPDGKIVYSSAL 345
QY 486 GAYDGKIKVMDLVAALDPAPAPACTLCLRTLVHSGRVFLQED 528
Db 346 GYQWDFASYDILRIKTPREVANLALL-----GFGDVFLALPD 383

RESULT 11

US-08-899-578-6
; Sequence 6, Application US/08899578
; Patent No. 6087153
; GENERAL INFORMATION:
; APPLICANT: Greenwald, Iva
; APPLICANT: Hubbard, E. Jane
; TITLE OF INVENTION: SEL-10 AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,578
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/53200/JPW/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 278-0525
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-899-578-6

Query Match 10.3%; Score 313.5; DB 3; Length 209;
Best Local Similarity 28.8%; Pred. No. 3.3e-25;
Matches 81; Conservative 37; Mismatches 80; Indels 83; Gaps 10;

QY 270 VYCLQYDDQKIYSGLRDNTIKIMDKNTLECKRIILNGHGSVYLCQYDE--RYIITGSSDS 327
Db 7 LTCMQHHDVLYVTGSDDNLTAKW-----CGHTGVTSTQISQCGRIYVSGSTDR 55
QY 328 TVRVMDVNTGEMINTLIHCEAVNLHLRFNNGMAMVTSKDRSLAVMDASPTDITLRLVLY 387
Db 56 TVKVMGHSITV-----RCMAA-----GSTLYVGSRTTILRVMD----- 90
QY 388 GIRAANVYDFPDKIYVSGSRTIKVNTSTCEVFTLNGHKRGACLOYDR--LVVS 445
Db 91 GHAAARCVQFPGTIVVSGGYDFYKIV-----NGHNRRYSLTFESERSIVCS 139
QY 446 GSSDNTIRLMDIEGACLRVLEGHFELVRCIRFDNKRIYSGAYDGKIKVMDLVAALDPRA 505
Db 140 GSLDTSIRVMD-----GHTSLTSGMQLRGNTILVSCNADSHVRVMD----- 179

QY 506 PAGTLCRLTVHSGRVFLQ--FDEQIVSSSHDITILMD 545
Db 180 -----GHRSAITSIQWFGRMVATSSDDGTVKLMD 209

RESULT 12

US-08-188-582-5
; Sequence 5, Application US/08188582
; Patent No. 5534410
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dylact, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; NUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,582
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RRO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 704 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-188-582-5

Query Match 10.1%; Score 306; DB 1; Length 704;
Best Local Similarity 22.8%; Pred. No. 1.6e-23;
Matches 122; Conservative 90; Mismatches 192; Indels 132; Gaps 22;

QY 95 EKEKELCVKFFQWSESDOVERVEVHLISQMC---HYQGHINSLY--KP--MLQRPFTT 146
Db 160 EKAKEFIEKY-----KCDLDGYYIGLFLNLLLSKPEELLENLDLV 200
QY 147 ALP-----ARGDHIENILSYLDAKSLCAELVCKEWTRYVSDGLMKLEIRVR 198
Db 201 AKEQDKFYIRMSRDSLSLKRHIQ--DRQEVYADIVSKYLFHFDYEGARKKL--QCYA 256
QY 199 TDSLWGLAERRGQOYLE-----KNRP----- 221
Db 257 TAGSHLGEAKRODNKMRVYGGLLKEVDFTLTTPAPAPEEDDDPDAPRPKKPKPKDP 316
QY 222 -----PDGNAPNPSFYALPKIITQDIETIESNMRCGRHSLORHGSE----- 265
Db 317 LLSKRSKSDPNAP--SIDRIPLPELKSDSLKL--LKALREASKRLALSQDLPSPAVFT 372

QY 266 ---TSKGVYCLQYDDOK--IVSGLRDNFTIKIMDKNTLECKRILNGHTGVSVCLOYDERVI 320
Db 373 VLNSHOGVTCALISDSDTMLACFGDSSVRIWSLTPAKRLTKDADS-----LREL 423
QY 321 ITGSSDSTVRVMDVNTGEMLTLLIHCBAVLHLP--NNGMAYTCSKDRSIAYVMDMASPT 378
Db 424 DKESADINVRMLDDRSRGVTRSLMGHTGPYRCAPAPENMLLSCSESDSTRILMSLTWS 483
QY 379 DITLRRVVGHRAAVNVDF--DDKIYASAGDRITIKYWNSTCEFPVTLNGHKGRIACL 436
Db 444 CVVYTR---GHVYPVMDVRFAPHGYYFVCSYDKTARLWATDSNALKEVGHLSDDCV 540
QY 437 QY--RDRLVVSGSSDNTIRLMDIEGACLRVLEGHHELVRCIRFD--NKRIYSGAYDKI 492
Db 541 QHPNSNYVATGSSDRYTRLMNDMTGOSVRLMTGHKGSVSLAFSACGRYIASSGVYDNI 600
QY 493 KVMDLVAALDPAPAGTLCRLTVEHSGRVFRLQF--DEFOIVSSSHDITLIMDF 546
Db 601 IIMDL-----SNGSL-VTTLRHTSTVTITFSRDGTVALAAGLDNNLTIMDF 647

RESULT 13

US-08-646-715-5
; Sequence 5, Application US/08646715
; Patent No. 5637686

GENERAL INFORMATION:

APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: DATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAPS AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-646-715-5

Query Match 10.1%; Score 306; DB 1; Length 704;
Best Local Similarity 22.8%; Pred. No. 1.6e-23;
Matches 122; Conservative 90; Mismatches 192; Indels 132; Gaps 22;

QY 95 EKEKELCVYFFEQMSRSDQVEFEVHLISQMC---HYOGHINSYL--KP--MLQDRFIT 146
Db 160 EKAKEFEIKY-----KCDLDGYIEGLFNLLLSKPEELIENDLVV 200
QY 147 ALP-----ARGLDHIAENLISYDAKSLCAAEYCKEWRYRSDGMKKLIERVYR 198
Db 201 AMEQDFVIRMSDSSHLEKRRHIQ--DRQEVYADIVSKYLFHFDYEGMARNKL--QCVA 256
QY 199 TDELMGLAEIRGQGYLF-----KNKP----- 221
Db 257 TAGSHLGEAKRQDNKKRIVYGLLEKVEYDPTLTTPAPAPEEDDDPADRPRKKKKRPP 316
QY 222 -----PDGNAPNSFYALYPKIITQDIETESNMWRCGRSLORICRSE----- 265
Db 317 LLSKKSKSDPNAP--SIDRIPLEKDSDKLK--LKALREASKRLALSKDQLPSAVFYT 372
QY 266 ---TSKGVYCLQYDDOK--IVSGLRDNFTIKIMDKNTLECKRILNGHTGVSVCLOYDERVI 320
Db 373 VLNSHOGVTCALISDSDTMLACFGDSSVRIWSLTPAKRLTKDADS-----LREL 423
QY 321 ITGSSDSTVRVMDVNTGEMLTLLIHCBAVLHLP--NNGMAYTCSKDRSIAYVMDMASPT 378
Db 424 DKESADINVRMLDDRSRGVTRSLMGHTGPYRCAPAPENMLLSCSESDSTRILMSLTWS 483
QY 379 DITLRRVVGHRAAVNVDF--DDKIYASAGDRITIKYWNSTCEFPVTLNGHKGRIACL 436
Db 484 CVVYTR---GHVYPVMDVRFAPHGYYFVCSYDKTARLWATDSNALKEVGHLSDDCV 540
QY 437 QY--RDRLVVSGSSDNTIRLMDIEGACLRVLEGHHELVRCIRFD--NKRIYSGAYDKI 492
Db 541 QHPNSNYVATGSSDRYTRLMNDMTGOSVRLMTGHKGSVSLAFSACGRYIASSGVYDNI 600
QY 493 KVMDLVAALDPAPAGTLCRLTVEHSGRVFRLQF--DEFOIVSSSHDITLIMDF 546
Db 601 IIMDL-----SNGSL-VTTLRHTSTVTITFSRDGTVALAAGLDNNLTIMDF 647

RESULT 14

US-08-190-802A-62
; Sequence 62, Application US/08190802A
; Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA

COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 704 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: TRANSCRIPTION FACTOR T1TF, Fig. 45
US-08-190-802A-62

Query Match 10.1%; Score 305; DB 1; Length 704;
Best Local Similarity 22.8%; Pred. No. 2,1e-23;
Matches 122; Conservative 89; Mismatches 193; Indels 132; Gaps 22;

QY 95 EKEKELCVKFEQWSESDQVEFVEHLISQMC---HYQGHINSYL--KP--MLQRFIT 146
DB 160 EKAKFEIKY-----KCDLDGYIEGLFNLLLSKPELLENDLVV 200
QY 147 ALP-----ARGLDHAENTLSYLDKSLCAAEVLCKEYRVTSDGMLKKLIERVYR 198
DB 201 AMEDKFEVIRMSRDSHSJFKRHQ--DRQEVVADIVSKYLHFDYEGMARNTL--QCVA 256
QY 199 TDSLWGLAERGMQGYLF-----KNKP----- 221
DB 257 TAGSHLGEAKKQDNKMRVYGLLEKVEVDFQTLTPAPAEEDDDPADRPRKKRPPKDP 316
QY 222 -----PDGNAPNSFYRALYPKTIIDETIESNMRCGRHSIQRHCRSE----- 265
DB 317 LLSKRSKSDPNAP--SIDRIPLEIKDSKDLK--LKALREASKRLALSKDQLPSAVEYT 372
QY 266 ---TSKGYVCQYDQK--IYSGLDNTIKITMDKNTLECKRLITGHSSVLCLOYDERVI 320
DB 373 VLNSHQGVTCAEISDSTMLACGFGDSSVRIWSTLPANVRLTKDADS-----LREL 423
QY 321 ITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRF--NNGMNVYTCSDKRSIAVWDMASPT 378
DB 424 DKESADIVRMILDDRSGEVTSLSMGHTGPYRCAPAPENMLLSSSESTIRLWLSLWTS 483
QY 379 DITLRRVLGHRAAVNVDF--DDKYIVASGDRITIKVWNTSTCEVFTLNGHKGIAQL 436
DB 484 CVVYTR--GHVYPMVDVRFAPHGYYFVSCSYDKFARLAMDSDNALRVFVGHLSDDVCV 540
QY 437 QY--RDRIVSGSSDNTIRLMDIEGACLRVLEGEHELVRICIRPD--NKRIYSGAYDGI 492
DB 541 QFHPSNTVATGSSDRTVRLMDNMTGOSVRLMTGKHGSVSSLAFCAGCRYLASGSVDHNI 600
QY 493 KVMDLVAALDPRAPAGTCLRTLVESHSGRVFRIQF--DEFQIVSSHDDTLILIMF 546
DB 601 IIMDL-----SNGSL-VTTLRHSTVTYTTITFSRDGYVLAAGLDNNLTIIMF 647

RESULT 15
US-08-308-818-3
Sequence 3, Application US/08308818
Patent No. 5847077
GENERAL INFORMATION:
APPLICANT: Green, Michael R
APPLICANT: Reese, Joseph C
TITLE OF INVENTION: A No. 5847077e1 Fungal Multisubunit Protein
TITLE OF INVENTION: Complex Critical for Expression of Fungal Proteins
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/308,818
FILING DATE: 19-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, S. Peter
REGISTRATION NUMBER: 25,351
REFERENCE/DOCKET NUMBER: 0342/0A404
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-52707700
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: D. melanogaster
IMMEDIATE SOURCE:
CLONE: TAFIT-80
US-08-308-818-3

Query Match 10.1%; Score 305; DB 2; Length 704;
Best Local Similarity 22.8%; Pred. No. 2,1e-23;
Matches 122; Conservative 89; Mismatches 193; Indels 132; Gaps 22;

QY 95 EKEKELCVKFEQWSESDQVEFVEHLISQMC---HYQGHINSYL--KP--MLQRFIT 146
DB 160 EKAKFEIKY-----KCDLDGYIEGLFNLLLSKPELLENDLVV 200
QY 147 ALP-----ARGLDHAENTLSYLDKSLCAAEVLCKEYRVTSDGMLKKLIERVYR 198
DB 201 AMEDKFEVIRMSRDSHSJFKRHQ--DRQEVVADIVSKYLHFDYEGMARNTL--QCVA 256
QY 199 TDSLWGLAERGMQGYLF-----KNKP----- 221
DB 257 TAGSHLGEAKKQDNKMRVYGLLEKVEVDFQTLTPAPAEEDDDPADRPRKKRPPKDP 316
QY 222 -----PDGNAPNSFYRALYPKTIIDETIESNMRCGRHSIQRHCRSE----- 265
DB 317 LLSKRSKSDPNAP--SIDRIPLEIKDSKDLK--LKALREASKRLALSKDQLPSAVEYT 372
QY 266 ---TSKGYVCQYDQK--IYSGLDNTIKITMDKNTLECKRLITGHSSVLCLOYDERVI 320
DB 373 VLNSHQGVTCAEISDSTMLACGFGDSSVRIWSTLPANVRLTKDADS-----LREL 423
QY 321 ITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRF--NNGMNVYTCSDKRSIAVWDMASPT 378
DB 424 DKESADIVRMILDDRSGEVTSLSMGHTGPYRCAPAPENMLLSSSESTIRLWLSLWTS 483
QY 379 DITLRRVLGHRAAVNVDF--DDKYIVASGDRITIKVWNTSTCEVFTLNGHKGIAQL 436
DB 484 CVVYTR--GHVYPMVDVRFAPHGYYFVSCSYDKFARLAMDSDNALRVFVGHLSDDVCV 540
QY 437 QY--RDRIVSGSSDNTIRLMDIEGACLRVLEGEHELVRICIRPD--NKRIYSGAYDGI 492
DB 541 QFHPSNTVATGSSDRTVRLMDNMTGOSVRLMTGKHGSVSSLAFCAGCRYLASGSVDHNI 600
QY 493 KVMDLVAALDPRAPAGTCLRTLVESHSGRVFRIQF--DEFQIVSSHDDTLILIMF 546
DB 601 IIMDL-----SNGSL-VTTLRHSTVTYTTITFSRDGYVLAAGLDNNLTIIMF 647

Thu Jun 7 10:44:14 2001

Search completed: June 7, 2001, 10:36:14
Job time: 73 sec

us-09-415-795-4.ra1

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 10:35:01 ; Search time 21.72 Seconds

(without alignments)
1800.338 Million cell updates/sec

Title: US-09-415-795-4

Perfect score: 3034

Sequence: 1 MDRAEAVLQEKALKFMSSE.....PAAQAEPPSPSRITYISR 569

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2597	85.6	518	BA8088	beta-transducin re
2	1635.5	53.9	701	T16607	hypothetical prote
3	690	22.7	506	T50211	WD-repeat protein
4	590.5	19.5	605	T38932	probable sulfur me
5	545	18.0	640	S49932	MET30 protein - ye
6	531.5	17.5	650	T46660	sulfur controller
7	520	17.1	579	T22703	hypothetical prote
8	519.5	17.1	267	S62507	hypothetical trp-a
9	455.5	15.0	1356	T18521	beta transducin-11
10	453	14.9	775	T45136	WD repeat protein
11	399	13.2	779	S56245	cell division cont
12	396.5	13.1	703	T43557	F-box/WD-repeat pr
13	354	11.7	409	S36113	LIS-1 protein - hu
14	354	11.7	410	S48052	platelet-activatin
15	341	11.2	515	T19487	hypothetical prote
16	339.5	11.2	777	T41075	hypothetical WD-re
17	337	11.1	1146	A55532	myosin-heavy-chain
18	332.5	11.0	317	T46032	WD-40 repeat regul
19	318	10.5	422	A56640	CDG4 repeat unit-c
20	313.5	10.3	376	T19266	hypothetical prote
21	312.5	10.3	714	S56893	hypothetical prote
22	312	10.3	876	T51507	WD40-repeat protei
23	307.5	10.1	1693	S76086	beta transducin-11
24	305	10.1	704	S33263	transcription init
25	299.5	9.9	1194	T03818	apoptotic proteina
26	297	9.8	502	T41148	trp-asp repeat con
27	292	9.6	659	S38108	hypothetical prote
28	290.5	9.6	614	S58305	WD-40 repeat regul
29	287.5	9.5	327	S48839	guanine nucleotide

30	285.5	9.4	798	2	S34023	TATA box-binding p
31	283	9.3	713	2	JN0133	WD-40 repeat regul
32	280	9.2	494	2	T19550	hypothetical prote
33	278.5	9.2	380	2	T40283	beta-transducin -
34	277	9.1	586	2	T38992	WD-40 repeat regul
35	276.5	9.1	501	2	T27513	hypothetical prote
36	274	9.0	2241	2	T16064	hypothetical prote
37	273	9.0	325	2	T09613	probable GTP-bindl
38	272.5	9.0	1008	2	T32986	hypothetical prote
39	271.5	8.9	817	2	S51445	probable membrane
40	270	8.9	451	2	S65162	hypothetical prote
41	270	8.9	473	2	T33805	hypothetical prote
42	270	8.9	504	2	T50983	probable pleiotrop
43	270	8.9	651	2	T50289	WD repeat protein
44	269.5	8.9	643	2	T41454	transcription init
45	265.5	8.8	325	2	T23309	hypothetical prote

ALIGNMENTS

RESULT 1

BA8088 beta-transducin repeat-containing protein - African clawed frog

N:Alternate names: beta-trcp

C:Species: Xenopus laevis (African clawed frog)

C>Date: 26-May-1994 #sequence #revision 26-May-1994 #text-change 21-Jul-2000

C:Accession: BA8088

R:Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A:Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase

A:Reference number: A48088; MUID:93330289

A:Accession: BA8088

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-518 <SP>

A:Cross-References: GB:M98268; NID:9295542; PIDN:AA02810.1; PID:9295543

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

C:Keywords: duplication

F:431-462/Domain: WD repeat homology <WD1>

Query Match	Score	2597	DB 2	Length	518
Best Local Similarity	91.6%	Pred.	No. 1.4e-193		
Matches	488	Conservative	7	Mismatches	8
				Indels	30
				Gaps	1
QY	18	SSPERDCNNGEPPRKTIIPKNSLRQYNSCARCLQNETVCLASTAMKTENCVAKTKIAN	77		
DB	13	ASEREDCNDEPPRKTIIPKNTLRO-----TKLAN	42		
QY	78	GTSSMIVPRQKLSAYEKEKEICVYFEQWSESDVEFEVHLISQMGYOHGHINSYLK	137		
DB	43	GTSSMIVPRQKLSAYEKEKEICVYFEQWSESDVEFEVHLISQMGYOHGHINSYLK	102		
QY	138	PMIQRDFITLAPRGIDHIAENILSYLDKSSCAELVCKEWMYRVTSQGMKKLIERAV	197		
DB	103	PMIQRDFITLAPRGIDHIAENILSYLDKSSCAELVCKEWMYRVTSQGMKKLIERAV	162		
QY	198	RHDSLMRGLAERGGQYLFKNKPPDGKTPNSFYKALPKIIOLETESMMRGGRHSL	257		
DB	163	RHDSLMRGLAERGGQYLFKNKPPDGKTPNSFYKALPKIIOLETESMMRGGRHSL	222		
QY	258	QRHCKSETSKGVYCLQYDQKIVSGIRDNITKIMDKNTECKRILITGHTSVLCQYDE	317		
DB	223	QRHCKSETSKGVYCLQYDQKIVSGIRDNITKIMDKNTECKRILITGHTSVLCQYDE	282		
QY	318	RYITIGSSDSYRVMDVNTGEMLNTLIHCEAVLHRENNGMATVCSKORSTAVWDMASP	377		
DB	283	RYITIGSSDSYRVMDVNTGEMLNTLIHCEAVLHRENNGMATVCSKORSTAVWDMASP	342		
QY	378	PTITLRRVLYGHRVAVNVVDPDKTYIVSASGRRTIKVWNTSCFPRTLNGHKGRIACIQ	437		
DB	343	PTITLRRVLYGHRVAVNVVDPDKTYIVSASGRRTIKVWNTSCFPRTLNGHKGRIACIQ	402		

Oy 438 YRDLRVGGSSNPTRLMDIEGACLRVLEGHEELVCRIRPDNRKRTVGAYDGKITWMDL 497
|||||
Db 403 YRDLRVGGSSNPTRLMDIEGACLRVLEGHEELVCRIRPDNRKRTVGAYDGKITWMDL 462

Oy 498 VVALDPRAPACTLCRTLVEHSGRFRFLQDFEFOIVSSSHDDTLIMPLNDP 550
|||||
Db 463 VVALDPRAPACTLCRTLVEHSGRFRFLQDFEFOIVSSSHDDTLIMPLNDP 515

RESULT 2
T16607
hypothetical protein K10B2.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16607
R:Miller, N.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid K10B2.
A:Reference number: Z18545
A:Accession: T16607
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-701 <MIL>
A:Cross-references: EMBL:U28730; NID:g860694; PID:g860695; PIDN:AAA68258.1; CESP:K10B2.1
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:K10B2.1
A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 666/3

```

Query Match 53.9%: Score 1635.5; DB 2; Length 701;
Best Local Similarity 57.7%: Pred. No. 7,6e-119;
Matches 328; Conservative 69; Mismatches 116; Indels 55; Gaps 8;

Oy 31 RRIPEKNSLQTYNSCARCLINDETVCLASTAMKTENCVAKTKLAN-----GTS 80
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 2 RRFEGKRALQKRGHARGGSGIALTVGVST---IERCF--TAVSNQIFELFTSFYSVF 55
Oy 81 SMIVKOR-----KLSAYEKKEKELCVKVFEDMSESDOVEFEVHLSOMCHYOHGHSY 135
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 56 SFLPFSRMTQFLFSTRFSFSFEVL-----KWSHEHEDLDENDKIVHRLSHYOLAKVNF 110
Oy 136 LKPLQDFTALPARGLDHAENILSYLAKSLCAELVCKEMYRVTSIDGLMKLLIER 195
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 111 IRPMLQRFISLTPA---HLVELILFVNSDSLKCEESTSMRCALARQGHMKLLIEK 166
Oy 196 MYRDTSLRGLAERGMGOYL-----FKNNPQGNAPNPSFYAL 235
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 167 NRSDSLWMLSEKKQMKPLNISIDMSVRRICEKENYDVNIKRDKLQLLIMHFYSKL 226
Oy 236 YPKIJDDETIESNMRGRHSRKHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMPN 295
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 227 YPKIRIDIHNDNNMKRGNYKMKTRINCQSENSKGYVCLQYDDOKIVSGLRDNTIKIMPR 286
Oy 296 TLECKRIITGHVGLCYQDERVITITGSSSDTVAMVPMVNGEMLNTLIHHCENLHLP 355
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 287 DYSGRILLSGHGVSLCYQDNRVITISGSDTVAMVPMVEGECKITLIHHCENLHLP 346
Oy 356 NNGMAYTCSKDRSLAVMDMASPTDITLRVYLGHRAAVNVYDFDQKIVTSASGDTIKW 415
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 347 ANGLMVTCSDKRSIAVMDMVSPTDITIRVYLGHRAAVNVYDFDQKIVTSASGDTIKW 406
Oy 416 NTSTCEPRTLNGHKGRIACLYQDRILVYSSSDNTILMLIEGACALRYLEGHELYRC 475
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 407 SMDLTLEPRTLAGHGRIGIACLYQGRILVYSSSDNTILMLIHSGVCALRYLEGHELYRC 466
Oy 476 IRPNKRIYSAVGDKIWMDLVALDPRAPAGTCLRTLVHSGRVRLQFDEQIYSS 535
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 467 IRPEKRIYSAVGDKIWMDLQALDPRALSELCLCSLVQHGVRVRLQFDDQIYSS 526
Oy 536 SHDITLIMDFLNDPAAQAEPPRSBRT 563
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Db 527 SHDDTILIMDFADP-----PSGLPST 549

RESULT 3

T50211

WD-repeat protein [Imported] - fission yeast (*Schizosaccharomyces pombe*)

C:Species: Schizosaccharomyces pombe

C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 02-Sep-2000

C:Accession: T50211

R:McDonally, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.;

Submitted to the EMBL Data Library, January 2000

A:Reference number: Z25046

A:Accession: T50211

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-506 <MCD>

A:Cross-references: EMBL:AL136538; PIDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05

A:Experimental source: strain 972h(-); cosmid c30

C:Genetics:

A:Gene: SPAC296E.01; SPDB:SPAC30.05

A:Map position: 1

A:Introns: 43/1; 74/3

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

[illegible]

[illegible]

```

QY 409 DRTVWNTSTCEPFRITLNGHKRGACLOVR--DRLVYSSGSDMTIRLMDCEGACLRVL 466
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1030 DRTIKWPIASTCTCTQTLLEGHGQVQSVSPDGGRAVSGSDDHTIKIMDVSGTCTQTL 1089
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 467 EGHLELVACICEF--DNKRIVSGAYDGKIKRWDLVAALDPAPAGTLCRLTVHSGRYFR 524
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1090 ESHGDSVSVAFSPDGGQVAVSGSIDGTIKIMDAAGTCTQ 1171
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 525 LQF--DEFQIVSSSHDITLIMDFLNDPAQ 553
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1141 VAFSPDGGQVAVSGSIDGTIKIMDAAGTCTQ 1171
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 10
T45136
WD repeat protein popl [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
C:Accession: T45136; T40157
R:Kohnami, K.; Toda, T.
submitted to the EMBL Data Library, September 1996
A:Description: Fission yeast WD repeat protein Popl is involved in maintenance of plo
A:Reference number: Z22925
A:Accession: T45136
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-775 <KOM>
A:Cross-references: EMBL:Y08391; PIDN:CAA69671.1
A:Experimental source: strain h- 972
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21842
A:Accession: T40157
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-775 <WOO>
A:Cross-references: EMBL:AL022103; PIDN:CA117898.1; GSPDB:GN00067; SPDB:SPBC262.18
A:Experimental source: strain 972h-; cosmid c262
C:Genetics:
A:Gene: SPBC262.18
A:Map position: 2
A:Note: popl+
Query Match 14.9%; Score 453; DB 2; Length 775;
Best Local Similarity 25.8%; Pred. No. 4.4e-27;
Matches 162; Conservative 86; Mismatches 229; Indels 150; Gaps 20;
QY 9 QEKALKFMNNSEREDC-----NNGEPPRKILPEKNSLRQTY-NSCARLQNET 56
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 REKCKLRNRSLSLNLHANKRFLFNSQDGKKNKTFPSTYVSNVFNNDCKREVAS 180
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 57 VC-----LASTAMKTENCVAKTYTLANGTSMIVPKORAKLSA----- 92
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 TFSLDAPNNVSVSYFSPNLLGNDKSTRQSPPHSSSHNSLHPVLYDSESNPSIHP 240
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 93 -----SYEKELK-----CYKYPEQWSESDQVEFVEHLISOMCHYQHGINSYLRKLPORDF 144
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 SNHLSQKNAVYLKLAQLISFEKLDPESYROYILFHLILSRGKHAVONHKKILPLFOKNF 300
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 145 ITALPARGLDIHAENILSYLDAKSLCAAEILYCKEYR-VTSDGMLMKL-----I 193
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 LTGFPA-----ETTNLVLHLDAPLSLCAVSOYSHMHYKLYVSSNEELMKSLPLKDFGFMDSI 356
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 194 EEMVVTDSIMRGLAERRGQYLFKNKPPDGNAPNNSFYRALYPIKTIIDTIESNRCG 253
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 357 DSKRTMCLQESLSA-----CAIMKRVYFRHNLREWR--- 389
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 254 RHSLDRHICRSEFTSGYVC---LOYDDQIVSGLDNFTIKIMDKWTECKRLLTGHTGSV 310
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 390 LHAPEKIRKCFPIHGVRLITKLPDFDDKIIIVTCSPTNINITYDKTGVLIINSLEHEG 449

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 10:36:01 ; Search time 13.01 Seconds

(without alignments)
1498.184 Million cell updates/sec

Title: US-09-415-795-4

Perfect score: 3034
Sequence: 1 MDPAAVLOEKALFKMNSSE.....PAAQAPPPSPRTTYISR 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3006	99.1	605	1	FW1A_HUMAN
2	2597	85.6	518	1	TRCB_XENLA
3	2384.5	78.6	542	1	FW1B_HUMAN
4	1635.5	53.9	701	1	YSS1_CAEEL
5	590.5	19.5	605	1	POB1_SCHPO
6	575	19.0	678	1	SCOB_EMENT
7	545	18.0	640	1	MT30_YEAST
8	531.5	17.5	650	1	SCQ2_NEUCR
9	520	17.1	579	1	SE10_CAEEL
10	519.5	17.1	267	1	YAF1_SCHPO
11	477.5	15.7	684	1	CC4_CANAL
12	455.5	15.0	1356	1	HEP1_PODAN
13	453	14.9	775	1	POB1_SCHPO
14	399	13.2	779	1	CC4_YEAST
15	396.5	13.1	703	1	POP2_SCHPO
16	392	12.9	732	1	KMBB_DICDI
17	354	11.7	409	1	LISI_HUMAN
18	354	11.7	409	1	LISI_MOUSE
19	353	11.6	409	1	LISI_BOVIN
20	341	11.2	515	1	YCM2_YEAST
21	337	11.1	1146	1	KMAA_DICDI
22	334.5	11.0	422	1	FMW2_HUMAN
23	325.5	10.7	361	1	WDS_DROME
24	324.5	10.7	362	1	WDS_HUMAN
25	318	10.5	422	1	FMW2_MOUSE
26	313.5	10.3	376	1	YKX4_CAEEL
27	312.5	10.3	714	1	YDL2_YEAST
28	307.5	10.1	742	1	PKWA_THECU
29	307.5	10.1	1693	1	Y163_SYNZ3
30	306	10.1	704	1	T2D4_DROME
31	299.5	9.9	1194	1	APAF_HUMAN
32	298.5	9.8	800	1	T2D4_HUMAN
33	292	9.6	659	1	YK16_YEAST

34	290.5	9.6	614	1	VA3A_SCHPO	Q09715 schizosacch
35	287.5	9.5	327	1	GLP1_BRANA	Q39336 brassica na
36	285.5	9.4	798	1	T2D4_YEAST	P38129 saccharomyc
37	283.5	9.3	327	1	GLP1_ARATH	Q24456 arabidopsis
38	283	9.3	473	1	PRL1_SCHPO	Q13615 schizosacch
39	283	9.3	713	1	TUPL1_YEAST	P16649 saccharomyc
40	282	9.3	318	1	GLP1_DROME	O18640 drosophila
41	281.5	9.3	682	1	TUPL1_KLULA	P56094 kluyveromyc
42	280.5	9.2	444	1	NDFE_EMENT	Q00664 emericella
43	276.5	9.1	501	1	YH92_CAEEL	Q23256 caenorhabdi
44	273	9.0	325	1	GLP1_MEDSA	Q24076 medicago sa
45	270	8.9	451	1	PRL1_YEAST	Q12417 saccharomyc

ALIGNMENTS

RESULT 1
FW1A_HUMAN STANDARD: PRT; 605 AA.
AC Q9Y297; Q9Y213;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP)
DE (E3RSIAPAPB) (PIKAPAPALPHA-E3 RECEPTOR SUBUNIT).
GN FBXW1A OR FBW1A OR BTRCP OR BTRC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=99075339; PubMed=9859996;
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RT Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;
RT "Identification of the receptor component of the IkkappaBalpha-
RT ubiquitin ligase.";
RL Nature 396:590-594(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Lymphoid;
RX MEDLINE=98325370; PubMed=9660940;
RA Margolin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,
RA Thomas D., Strebel K., Benarous R.;
RT "A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu
RT connects CD4 to the ER degradation pathway through an F-box motif.";
RL Mol. Cell 1:565-574(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Cenciarelli C., Chiau D.S., Guardavaccaro D., Parks W., Vidal M.,
RA Pagano M.;
RT "Identification of a family of human F-box proteins.";
RL Curr. Biol. 9:1177-1179(1999).
RN [4]
RP CHARACTERIZATION.
RA Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,
RA Harper J.W.;
RT "The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically
RT with phosphorylated destruction motifs in I-kappa-B-alpha and
RT beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro.";
RL Genes Dev. 13:270-283(1999).
CC -I- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA
CC (PIKAPAPALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR
CC UBIQUITINATION AND DEGRADATION.
CC -I- SUBUNIT: PART OF A SCF (SKP1-CULIN-F-BOX) PROTEIN LIGASE COMPLEX.
CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -I- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -I- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF101784; AAD08702.1; -
 DR EMBL: Y14153; CAA74572.1; -
 DR EMBL: AF129530; AAF04464.1; -
 DR MIM: 603482; -
 DR InterPro: IPR001680; -
 DR InterPro: IPR001810; -
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KM Ubiquitin conjugation: Repeat: WD repeat; Alternative splicing.
 FT DOMAIN 190 228 F-box.
 FT REPEAT 301 338 WD 1.
 FT REPEAT 341 378 WD 2.
 FT REPEAT 381 418 WD 3.
 FT REPEAT 424 461 WD 4.
 FT REPEAT 464 503 WD 5.
 FT REPEAT 505 541 WD 6.
 FT REPEAT 553 590 WD 7.
 FT VARSLIC 11 52 MISSING (IN ISIFORM 2).
 SQ SEQUENCE 605 AA; 68866 MW; 4C67F3B7E400FD37 CRC64;

Query Match 99.1%; Score 3006; DB 1: Length 605;
 Best Local Similarity 94.0%; Pred. No. 1.4e-215;
 Matches 369; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MDAEAVLDEKALFKM-----NSSREDC 24
 DB 1 MDAEAVLDEKALFKMSPRSLIMLGSSSLADSMPSRLCLYNGTGALAFQNSSREDC 60
 QY 25 NNGEPPRKIIPEKNSLRQTYNSCARCLNOETVCLASTAKTENCYAKTKLANGTSMIV 84
 DB 61 NNGEPPRKIIPEKNSLRQTYNSCARCLNOETVCLASTAKTENCYAKTKLANGTSMIV 120
 QY 85 PKRKLSASERKEKLVYFEQWSESDVEFEHLISQCHYOHGHINSYKPMQRDP 144
 DB 121 PKRKLSASERKEKLVYFEQWSESDVEFEHLISQCHYOHGHINSYKPMQRDP 180
 QY 145 ITPALPARGDHIADENILSYLDAKSLCAELVCKEYRVTSDDGLMKKLIERMVRTDSLWR 204
 DB 181 ITPALPARGDHIADENILSYLDAKSLCAELVCKEYRVTSDDGLMKKLIERMVRTDSLWR 240
 QY 205 GLAERRGWGYLFKNRPDGNAPNSFYALYKIIODIETTESNRCGRHSLSQRIHCS 264
 DB 241 GLAERRGWGYLFKNRPDGNAPNSFYALYKIIODIETTESNRCGRHSLSQRIHCS 300
 QY 265 ETSKGYVCLQYDDOKIVSGLRONTIKIMPKNTECKRIITGHGYSVLCQYDERVITGS 324
 DB 301 ETSKGYVCLQYDDOKIVSGLRONTIKIMPKNTECKRIITGHGYSVLCQYDERVITGS 360
 QY 325 SDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGAMVYTSKDRSLAVWDMASPTDITLRR 384
 DB 361 SDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGAMVYTSKDRSLAVWDMASPTDITLRR 420
 QY 385 VLVGHRAAVNVDFDKYIVYASGDRITIVWNTSTEEVRYTLNGHKGRIACIQYRDLVY 444
 DB 421 VLVGHRAAVNVDFDKYIVYASGDRITIVWNTSTEEVRYTLNGHKGRIACIQYRDLVY 480
 QY 445 SGSSDNTIRLMDIEGACRLVLEGHEILVRCIRFDNKRIVSGAYDGKIRVMDLVALLDR 504
 DB 481 SGSSDNTIRLMDIEGACRLVLEGHEILVRCIRFDNKRIVSGAYDGKIRVMDLVALLDR 540

QY 505 APAGTLCRTIVHSGRVRLQFDFEQIVSSSHDITLIMDFLNDPAQAQEPSPSPRTY 564
 DB 541 APAGTLCRTIVHSGRVRLQFDFEQIVSSSHDITLIMDFLNDPAQAQEPSPSPRTY 600
 QY 565 TYISR 569
 DB 601 TYISR 605

RESULT 2

TRCB_XENLA STANDARD: PRT: 518 AA.
 AC Q91854; P70037; P70038;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE BETA-TRCP (BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN).
 GN FBXW1 OR BTCP.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RX MEDLINE=93330289; PubMed=8393141;
 RA Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
 RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
 RT anaphase are rescued by Xenopus CDNs encoding N-ras or a protein
 RT with beta-transducin repeats.";
 RL Mol. Cell. Biol. 13:4953-4966(1993).
 RN [2]
 RP SEQUENCE OF 302-518 FROM N.A.
 RA MEDLINE=97109804; PubMed=8952061;
 RT Hudson J.W., Alarcon V.B., Elinson R.P.;
 RT "Identification of new localized RNAs in the Xenopus oocyte by
 RT differential display PCR.";
 RL Dev. Genet. 19:190-198(1996).
 CC -1- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PHOSPHORYLATED
 CC PROTEINS AND PROMOTES THEIR UBIQUITINATION AND DEGRADATION.
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
 CC (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
 CC MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
 CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
 CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBD, OR
 CC TADPOLE EMBRYO.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: M98268; AAA02810.1; -
 EMBL: U63921; AAB49671.1; -
 EMBL: U63922; AAB49672.1; -
 InterPro: IPR001680; -
 InterPro: IPR001810; -
 Pfam: PF00646; F-box; 1.
 Pfam: PF00400; WD40; 7.
 PRINTS: PR00320; GPROTEINRPT.
 PROSITE: PS50181; FBOX; 1.
 PROSITE: PS00678; WD_REPEATS_1; 6.
 PROSITE: PS50082; WD_REPEATS_2; 7.
 PROSITE: PS50294; WD_REPEATS_REGION; 1.
 Ubiquitin conjugation: Repeat: WD repeat.
 DOMAIN 119 157 F-box.

```

FT REPEAT 230 258 WD 1.
FT REPEAT 270 298 WD 2.
FT REPEAT 310 338 WD 3.
FT REPEAT 353 381 WD 4.
FT REPEAT 393 421 WD 5.
FT REPEAT 433 461 WD 6.
FT REPEAT 482 510 WD 7.
FT REPEAT 516 548 WD 8.
FT REPEAT 582 610 WD 9.
FT REPEAT 616 644 WD 10.
FT REPEAT 650 678 WD 11.
FT REPEAT 684 712 WD 12.
FT REPEAT 718 746 WD 13.
FT REPEAT 752 780 WD 14.
FT REPEAT 786 814 WD 15.
FT REPEAT 820 848 WD 16.
FT REPEAT 854 882 WD 17.
FT REPEAT 888 916 WD 18.
FT REPEAT 920 948 WD 19.
FT REPEAT 954 982 WD 20.
FT REPEAT 986 1014 WD 21.
FT REPEAT 1020 1048 WD 22.
FT REPEAT 1062 1090 WD 23.
FT REPEAT 1094 1122 WD 24.
FT REPEAT 1136 1164 WD 25.
FT REPEAT 1178 1206 WD 26.
FT REPEAT 1210 1238 WD 27.
FT REPEAT 1242 1270 WD 28.
FT REPEAT 1274 1302 WD 29.
FT REPEAT 1306 1334 WD 30.
FT REPEAT 1336 1364 WD 31.
FT REPEAT 1378 1406 WD 32.
FT REPEAT 1410 1438 WD 33.
FT REPEAT 1450 1478 WD 34.
FT REPEAT 1490 1518 WD 35.
FT REPEAT 1530 1558 WD 36.
FT REPEAT 1570 1598 WD 37.
FT REPEAT 1610 1638 WD 38.
FT REPEAT 1650 1678 WD 39.
FT REPEAT 1690 1718 WD 40.
FT REPEAT 1730 1758 WD 41.
FT REPEAT 1770 1798 WD 42.
FT REPEAT 1810 1838 WD 43.
FT REPEAT 1850 1878 WD 44.
FT REPEAT 1890 1918 WD 45.
FT REPEAT 1930 1958 WD 46.
FT REPEAT 1970 1998 WD 47.
FT REPEAT 2010 2038 WD 48.
FT REPEAT 2050 2078 WD 49.
FT REPEAT 2090 2118 WD 50.
FT REPEAT 2130 2158 WD 51.
FT REPEAT 2170 2198 WD 52.
FT REPEAT 2210 2238 WD 53.
FT REPEAT 2250 2278 WD 54.
FT REPEAT 2290 2318 WD 55.
FT REPEAT 2330 2358 WD 56.
FT REPEAT 2370 2398 WD 57.
FT REPEAT 2410 2438 WD 58.
FT REPEAT 2450 2478 WD 59.
FT REPEAT 2490 2518 WD 60.
FT REPEAT 2530 2558 WD 61.
FT REPEAT 2570 2598 WD 62.
FT REPEAT 2610 2638 WD 63.
FT REPEAT 2650 2678 WD 64.
FT REPEAT 2690 2718 WD 65.
FT REPEAT 2730 2758 WD 66.
FT REPEAT 2770 2798 WD 67.
FT REPEAT 2810 2838 WD 68.
FT REPEAT 2850 2878 WD 69.
FT REPEAT 2890 2918 WD 70.
FT REPEAT 2930 2958 WD 71.
FT REPEAT 2970 2998 WD 72.
FT REPEAT 3010 3038 WD 73.
FT REPEAT 3050 3078 WD 74.
FT REPEAT 3090 3118 WD 75.
FT REPEAT 3130 3158 WD 76.
FT REPEAT 3170 3198 WD 77.
FT REPEAT 3210 3238 WD 78.
FT REPEAT 3250 3278 WD 79.
FT REPEAT 3290 3318 WD 80.
FT REPEAT 3330 3358 WD 81.
FT REPEAT 3370 3398 WD 82.
FT REPEAT 3410 3438 WD 83.
FT REPEAT 3450 3478 WD 84.
FT REPEAT 3490 3518 WD 85.
FT REPEAT 3530 3558 WD 86.
FT REPEAT 3570 3598 WD 87.
FT REPEAT 3610 3638 WD 88.
FT REPEAT 3650 3678 WD 89.
FT REPEAT 3690 3718 WD 90.
FT REPEAT 3730 3758 WD 91.
FT REPEAT 3770 3798 WD 92.
FT REPEAT 3810 3838 WD 93.
FT REPEAT 3850 3878 WD 94.
FT REPEAT 3890 3918 WD 95.
FT REPEAT 3930 3958 WD 96.
FT REPEAT 3970 3998 WD 97.
FT REPEAT 4010 4038 WD 98.
FT REPEAT 4050 4078 WD 99.
FT REPEAT 4090 4118 WD 100.
FT REPEAT 4130 4158 WD 101.
FT REPEAT 4170 4198 WD 102.
FT REPEAT 4210 4238 WD 103.
FT REPEAT 4250 4278 WD 104.
FT REPEAT 4290 4318 WD 105.
FT REPEAT 4330 4358 WD 106.
FT REPEAT 4370 4398 WD 107.
FT REPEAT 4410 4438 WD 108.
FT REPEAT 4450 4478 WD 109.
FT REPEAT 4490 4518 WD 110.
FT REPEAT 4530 4558 WD 111.
FT REPEAT 4570 4598 WD 112.
FT REPEAT 4610 4638 WD 113.
FT REPEAT 4650 4678 WD 114.
FT REPEAT 4690 4718 WD 115.
FT REPEAT 4730 4758 WD 116.
FT REPEAT 4770 4798 WD 117.
FT REPEAT 4810 4838 WD 118.
FT REPEAT 4850 4878 WD 119.
FT REPEAT 4890 4918 WD 120.
FT REPEAT 4930 4958 WD 121.
FT REPEAT 4970 4998 WD 122.
FT REPEAT 5010 5038 WD 123.
FT REPEAT 5050 5078 WD 124.
FT REPEAT 5090 5118 WD 125.
FT REPEAT 5130 5158 WD 126.
FT REPEAT 5170 5198 WD 127.
FT REPEAT 5210 5238 WD 128.
FT REPEAT 5250 5278 WD 129.
FT REPEAT 5290 5318 WD 130.
FT REPEAT 5330 5358 WD 131.
FT REPEAT 5370 5398 WD 132.
FT REPEAT 5410 5438 WD 133.
FT REPEAT 5450 5478 WD 134.
FT REPEAT 5490 5518 WD 135.
FT REPEAT 5530 5558 WD 136.
FT REPEAT 5570 5598 WD 137.
FT REPEAT 5610 5638 WD 138.
FT REPEAT 5650 5678 WD 139.
FT REPEAT 5690 5718 WD 140.
FT REPEAT 5730 5758 WD 141.
FT REPEAT 5770 5798 WD 142.
FT REPEAT 5810 5838 WD 143.
FT REPEAT 5850 5878 WD 144.
FT REPEAT 5890 5918 WD 145.
FT REPEAT 5930 5958 WD 146.
FT REPEAT 5970 5998 WD 147.
FT REPEAT 6010 6038 WD 148.
FT REPEAT 6050 6078 WD 149.
FT REPEAT 6090 6118 WD 150.
FT REPEAT 6130 6158 WD 151.
FT REPEAT 6170 6198 WD 152.
FT REPEAT 6210 6238 WD 153.
FT REPEAT 6250 6278 WD 154.
FT REPEAT 6290 6318 WD 155.
FT REPEAT 6330 6358 WD 156.
FT REPEAT 6370 6398 WD 157.
FT REPEAT 6410 6438 WD 158.
FT REPEAT 6450 6478 WD 159.
FT REPEAT 6490 6518 WD 160.
FT REPEAT 6530 6558 WD 161.
FT REPEAT 6570 6598 WD 162.
FT REPEAT 6610 6638 WD 163.
FT REPEAT 6650 6678 WD 164.
FT REPEAT 6690 6718 WD 165.
FT REPEAT 6730 6758 WD 166.
FT REPEAT 6770 6798 WD 167.
FT REPEAT 6810 6838 WD 168.
FT REPEAT 6850 6878 WD 169.
FT REPEAT 6890 6918 WD 170.
FT REPEAT 6930 6958 WD 171.
FT REPEAT 6970 6998 WD 172.
FT REPEAT 7010 7038 WD 173.
FT REPEAT 7050 7078 WD 174.
FT REPEAT 7090 7118 WD 175.
FT REPEAT 7130 7158 WD 176.
FT REPEAT 7170 7198 WD 177.
FT REPEAT 7210 7238 WD 178.
FT REPEAT 7250 7278 WD 179.
FT REPEAT 7290 7318 WD 180.
FT REPEAT 7330 7358 WD 181.
FT REPEAT 7370 7398 WD 182.
FT REPEAT 7410 7438 WD 183.
FT REPEAT 7450 7478 WD 184.
FT REPEAT 7490 7518 WD 185.
FT REPEAT 7530 7558 WD 186.
FT REPEAT 7570 7598 WD 187.
FT REPEAT 7610 7638 WD 188.
FT REPEAT 7650 7678 WD 189.
FT REPEAT 7690 7718 WD 190.
FT REPEAT 7730 7758 WD 191.
FT REPEAT 7770 7798 WD 192.
FT REPEAT 7810 7838 WD 193.
FT REPEAT 7850 7878 WD 194.
FT REPEAT 7890 7918 WD 195.
FT REPEAT 7930 7958 WD 196.
FT REPEAT 7970 7998 WD 197.
FT REPEAT 8010 8038 WD 198.
FT REPEAT 8050 8078 WD 199.
FT REPEAT 8090 8118 WD 200.
FT REPEAT 8130 8158 WD 201.
FT REPEAT 8170 8198 WD 202.
FT REPEAT 8210 8238 WD 203.
FT REPEAT 8250 8278 WD 204.
FT REPEAT 8290 8318 WD 205.
FT REPEAT 8330 8358 WD 206.
FT REPEAT 8370 8398 WD 207.
FT REPEAT 8410 8438 WD 208.
FT REPEAT 8450 8478 WD 209.
FT REPEAT 8490 8518 WD 210.
FT REPEAT 8530 8558 WD 211.
FT REPEAT 8570 8598 WD 212.
FT REPEAT 8610 8638 WD 213.
FT REPEAT 8650 8678 WD 214.
FT REPEAT 8690 8718 WD 215.
FT REPEAT 8730 8758 WD 216.
FT REPEAT 8770 8798 WD 217.
FT REPEAT 8810 8838 WD 218.
FT REPEAT 8850 8878 WD 219.
FT REPEAT 8890 8918 WD 220.
FT REPEAT 8930 8958 WD 221.
FT REPEAT 8970 8998 WD 222.
FT REPEAT 9010 9038 WD 223.
FT REPEAT 9050 9078 WD 224.
FT REPEAT 9090 9118 WD 225.
FT REPEAT 9130 9158 WD 226.
FT REPEAT 9170 9198 WD 227.
FT REPEAT 9210 9238 WD 228.
FT REPEAT 9250 9278 WD 229.
FT REPEAT 9290 9318 WD 230.
FT REPEAT 9330 9358 WD 231.
FT REPEAT 9370 9398 WD 232.
FT REPEAT 9410 9438 WD 233.
FT REPEAT 9450 9478 WD 234.
FT REPEAT 9490 9518 WD 235.
FT REPEAT 9530 9558 WD 236.
FT REPEAT 9570 9598 WD 237.
FT REPEAT 9610 9638 WD 238.
FT REPEAT 9650 9678 WD 239.
FT REPEAT 9690 9718 WD 240.
FT REPEAT 9730 9758 WD 241.
FT REPEAT 9770 9798 WD 242.
FT REPEAT 9810 9838 WD 243.
FT REPEAT 9850 9878 WD 244.
FT REPEAT 9890 9918 WD 245.
FT REPEAT 9930 9958 WD 246.
FT REPEAT 9970 9998 WD 247.
FT REPEAT 10010 10038 WD 248.
FT REPEAT 10050 10078 WD 249.
FT REPEAT 10090 10118 WD 250.
FT REPEAT 10130 10158 WD 251.
FT REPEAT 10170 10198 WD 252.
FT REPEAT 10210 10238 WD 253.
FT REPEAT 10250 10278 WD 254.
FT REPEAT 10290 10318 WD 255.
FT REPEAT 10330 10358 WD 256.
FT REPEAT 10370 10398 WD 257.
FT REPEAT 10410 10438 WD 258.
FT REPEAT 10450 10478 WD 259.
FT REPEAT 10490 10518 WD 260.
FT REPEAT 10530 10558 WD 261.
FT REPEAT 10570 10598 WD 262.
FT REPEAT 10610 10638 WD 263.
FT REPEAT 10650 10678 WD 264.
FT REPEAT 10690 10718 WD 265.
FT REPEAT 10730 10758 WD 266.
FT REPEAT 10770 10798 WD 267.
FT REPEAT 10810 10838 WD 268.
FT REPEAT 10850 10878 WD 269.
FT REPEAT 10890 10918 WD 270.
FT REPEAT 10930 10958 WD 271.
FT REPEAT 10970 10998 WD 272.
FT REPEAT 11010 11038 WD 273.
FT REPEAT 11050 11078 WD 274.
FT REPEAT 11090 11118 WD 275.
FT REPEAT 11130 11158 WD 276.
FT REPEAT 11170 11198 WD 277.
FT REPEAT 11210 11238 WD 278.
FT REPEAT 11250 11278 WD 279.
FT REPEAT 11290 11318 WD 280.
FT REPEAT 11330 11358 WD 281.
FT REPEAT 11370 11398 WD 282.
FT REPEAT 11410 11438 WD 283.
FT REPEAT 11450 11478 WD 284.
FT REPEAT 11490 11518 WD 285.
FT REPEAT 11530 11558 WD 286.
FT REPEAT 11570 11598 WD 287.
FT REPEAT 11610 11638 WD 288.
FT REPEAT 11650 11678 WD 289.
FT REPEAT 11690 11718 WD 290.
FT REPEAT 11730 11758 WD 291.
FT REPEAT 11770 11798 WD 292.
FT REPEAT 11810 11838 WD 293.
FT REPEAT 11850 11878 WD 294.
FT REPEAT 11890 11918 WD 295.
FT REPEAT 11930 11958 WD 296.
FT REPEAT 11970 11998 WD 297.
FT REPEAT 12010 12038 WD 298.
FT REPEAT 12050 12078 WD 299.
FT REPEAT 12090 12118 WD 300.
FT REPEAT 12130 12158 WD 301.
FT REPEAT 12170 12198 WD 302.
FT REPEAT 12210 12238 WD 303.
FT REPEAT 12250 12278 WD 304.
FT REPEAT 12290 12318 WD 305.
FT REPEAT 12330 12358 WD 306.
FT REPEAT 12370 12398 WD 307.
FT REPEAT 12410 12438 WD 308.
FT REPEAT 12450 12478 WD 309.
FT REPEAT 12490 12518 WD 310.
FT REPEAT 12530 12558 WD 311.
FT REPEAT 12570 12598 WD 312.
FT REPEAT 12610 12638 WD 313.
FT REPEAT 12650 12678 WD 314.
FT REPEAT 12690 12718 WD 315.
FT REPEAT 12730 12758 WD 316.
FT REPEAT 12770 12798 WD 317.
FT REPEAT 12810 12838 WD 318.
FT REPEAT 12850 12878 WD 319.
FT REPEAT 12890 12918 WD 320.
FT REPEAT 12930 12958 WD 321.
FT REPEAT 12970 12998 WD 322.
FT REPEAT 13010 13038 WD 323.
FT REPEAT 13050 13078 WD 324.
FT REPEAT 13090 13118 WD 325.
FT REPEAT 13130 13158 WD 326.
FT REPEAT 13170 13198 WD 327.
FT REPEAT 13210 13238 WD 328.
FT REPEAT 13250 13278 WD 329.
FT REPEAT 13290 13318 WD 330.
FT REPEAT 13330 13358 WD 331.
FT REPEAT 13370 13398 WD 332.
FT REPEAT 13410 13438 WD 333.
FT REPEAT 13450 13478 WD 334.
FT REPEAT 13490 13518 WD 335.
FT REPEAT 13530 13558 WD 336.
FT REPEAT 13570 13598 WD 337.
FT REPEAT 13610 13638 WD 338.
FT REPEAT 13650 13678 WD 339.
FT REPEAT 13690 13718 WD 340.
FT REPEAT 13730 13758 WD 341.
FT REPEAT 13770 13798 WD 342.
FT REPEAT 13810 13838 WD 343.
FT REPEAT 13850 13878 WD 344.
FT REPEAT 13890 13918 WD 345.
FT REPEAT 13930 13958 WD 346.
FT REPEAT 13970 13998 WD 347.
FT REPEAT 14010 14038 WD 348.
FT REPEAT 14050 14078 WD 349.
FT REPEAT 14090 14118 WD 350.
FT REPEAT 14130 14158 WD 351.
FT REPEAT 14170 14198 WD 352.
FT REPEAT 14210 14238 WD 353.
FT REPEAT 14250 14278 WD 354.
FT REPEAT 14290 14318 WD 355.
FT REPEAT 14330 14358 WD 356.
FT REPEAT 14370 14398 WD 357.
FT REPEAT 14410 14438 WD 358.
FT REPEAT 14450 14478 WD 359.
FT REPEAT 14490 14518 WD 360.
FT REPEAT 14530 14558 WD 361.
FT REPEAT 14570 14598 WD 362.
FT REPEAT 14610 14638 WD 363.
FT REPEAT 14650 14678 WD 364.
FT REPEAT 14690 14718 WD 365.
FT REPEAT 14730 14758 WD 366.
FT REPEAT 14770 14798 WD 367.
FT REPEAT 14810 14838 WD 368.
FT REPEAT 14850 14878 WD 369.
FT REPEAT 14890 14918 WD 370.
FT REPEAT 14930 14958 WD 371.
FT REPEAT 14970 14998 WD 372.
FT REPEAT 15010 15038 WD 373.
FT REPEAT 15050 15078 WD 374.
FT REPEAT 15090 15118 WD 375.
FT REPEAT 15130 15158 WD 376.
FT REPEAT 15170 15198 WD 377.
FT REPEAT 15210 15238 WD 378.
FT REPEAT 15250 15278 WD 379.
FT REPEAT 15290 15318 WD 380.
FT REPEAT 15330 15358 WD 381.
FT REPEAT 15370 15398 WD 382.
FT REPEAT 15410 15438 WD 383.
FT REPEAT 15450 15478 WD 384.
FT REPEAT 15490 15518 WD 385.
FT REPEAT 15530 15558 WD 386.
FT REPEAT 15570 15598 WD 387.
FT REPEAT 15610 15638 WD 388.
FT REPEAT 15650 15678 WD 389.
FT REPEAT 15690 15718 WD 390.
FT REPEAT 15730 15758 WD 391.
FT REPEAT 15770 15798 WD 392.
FT REPEAT 15810 15838 WD 393.
FT REPEAT 15850 15878 WD 394.
FT REPEAT 15890 15918 WD 395.
FT REPEAT 15930 15958 WD 396.
FT REPEAT 15970 15998 WD 397.
FT REPEAT 16010 16038 WD 398.
FT REPEAT 16050 16078 WD 399.
FT REPEAT 16090 16118 WD 400.
FT REPEAT 16130 16158 WD 401.
FT REPEAT 16170 16198 WD 402.
FT REPEAT 16210 16238 WD 403.
FT REPEAT 16250 16278 WD 404.
FT REPEAT 16290 16318 WD 405.
FT REPEAT 16330 16358 WD 406.
FT REPEAT 16370 16398 WD 407.
FT REPEAT 16410 16438 WD 408.
FT REPEAT 16450 16478 WD 409.
FT REPEAT 16490 16518 WD 410.
FT REPEAT 16530 16558 WD 411.
FT REPEAT 16570 16598 WD 412.
FT REPEAT 16610 16638 WD 413.
FT REPEAT 16650 16678 WD 414.
FT REPEAT 16690 16718 WD 415.
FT REPEAT 16730 16758 WD 416.
FT REPEAT 16770 16798 WD 417.
FT REPEAT 16810 16838 WD 418.
FT REPEAT 16850 16878 WD 419.
FT REPEAT 16890 16918 WD 420.
FT REPEAT 16930 16958 WD 421.
FT REPEAT 16970 16998 WD 422.
FT REPEAT 17010 17038 WD 423.
FT REPEAT 17050 17078 WD 424.
FT REPEAT 17090 17118 WD 425.
FT REPEAT 17130 17158 WD 426.
FT REPEAT 17170 17198 WD 427.
FT REPEAT 17210 17238 WD 428.
FT REPEAT 17250 17278 WD 429.
FT REPEAT 17290 17318 WD 430.
FT REPEAT 17330 17358 WD 431.
FT REPEAT 17370 17398 WD 432.
FT REPEAT 17410 17438 WD 433.
FT REPEAT 17450 17478 WD 434.
FT REPEAT 17490 17518 WD 435.
FT REPEAT 17530 17558 WD 436.
FT REPEAT 17570 17598 WD 437.
FT REPEAT 17610 17638 WD 438.
FT REPEAT 17650 17678 WD 439.
FT REPEAT 17690 17718 WD 440.
FT REPEAT 17730 17758 WD 441.
FT REPEAT 17770 17798 WD 442.
FT REPEAT 17810 17838 WD 443.
FT REPEAT 17850 17878 WD 444.
FT REPEAT 17890 17918 WD 445.
FT REPEAT 17930 17958 WD 446.
FT REPEAT 17970 17998 WD 447.
FT REPEAT 18010 18038 WD 448.
FT REPEAT 18050 18078 WD 449.
FT REPEAT 18090 18118 WD 450.
FT REPEAT 18130 18158 WD 451.
FT REPEAT 18170 18198 WD 452.
FT REPEAT 18210 18238 WD 453.
FT REPEAT 18250 18278 WD 454.
FT REPEAT 18290 18318 WD 455.
FT REPEAT 18330 18358 WD 456.
FT REPEAT 18370 18398 WD 457.
FT REPEAT 18410 18438 WD 458.
FT REPEAT 18450 18478 WD 459.
FT REPEAT 18490 18518 WD 460.
FT REPEAT 18530 18558 WD 461.
FT REPEAT 18570 18598 WD 462.
FT REPEAT 18610 18638 WD 463.
FT REPEAT 18650 18678 WD 464.
FT REPEAT 18690 18718 WD 465.
FT REPEAT 18730 18758 WD 466.
FT REPEAT 18770 18798 WD 467.
FT REPEAT 18810 18838 WD 468.
FT REPEAT 18850 18878 WD 469.
FT REPEAT 18890 18918 WD 470.
FT REPEAT 18930 18958 WD 471.
FT REPEAT 18970 18998 WD 472.
FT REPEAT 19010 19038 WD 473.
FT REPEAT 19050 19078 WD 474.
FT REPEAT 19090 19118 WD 475.
FT REPEAT 19130 19158 WD 476.
FT REPEAT 19170 19198 WD 477.
FT REPEAT 19210 19238 WD 478.
FT REPEAT 19250 19278 WD 479.
FT REPEAT 19290 19318 WD 480.
FT REPEAT 19330 19358 WD 481.
FT REPEAT 19370 19398 WD 482.
FT REPEAT 19410 19438 WD 483.
FT REPEAT 19450 19478 WD 484.
FT REPEAT 19490 19518 WD 485.
FT REPEAT 19530 19558 WD 486.
FT REPEAT 19570 19598 WD 487.
FT REPEAT 19610 19638 WD 488.
FT REPEAT 19650 19678 WD 489.
FT REPEAT 19690 19718 WD 490.
FT REPEAT 19730 19758 WD 491.
FT REPEAT 19770 19798 WD 492.
FT REPEAT 19810 19838 WD 493.
FT REPEAT 19850 19878 WD 494.
FT REPEAT 19890 19918 WD 495.
FT REPEAT 19930 19958 WD 496.
FT REPEAT 19970 19998 WD 497.
FT REPEAT 20010 20038 WD 498.
FT REPEAT 20050 20078 WD 499.
FT REPEAT 20090 20118 WD 500.
FT REPEAT 20130 20158 WD 501.
FT REPEAT 20170 20198 WD 502.
FT REPEAT 20210 20238 WD 503.
FT REPEAT 20250 20278 WD 504.
FT REPEAT 20290 20318 WD 505.
FT REPEAT 20330 20358 WD 506.
FT REPEAT 20370 20398 WD 507.
FT REPEAT 20410 20438 WD 508.
FT REPEAT 20450 20478 WD 509.
FT REPEAT 20490 20518 WD 510.
FT REPEAT 20530 20558 WD 511.
FT REPEAT 20570 20598 WD 512.
FT REPEAT 20610 20638 WD 513.
FT REPEAT 20650 20678 WD 514.
FT REPEAT 20690 20718 WD 515.
FT REPEAT 20730 20758 WD 516.
FT REPEAT 20770 20798 WD 517.
FT REPEAT 20810 20838 WD 518.
FT REPEAT 20850 20878 WD 519.
FT REPEAT 20890 20918 WD 520.
FT REPEAT 20930 20958 WD 521.
FT REPEAT 20970 20998 WD 522.
FT REPEAT 21010 21038 WD 523.
FT REPEAT 21050 21078 WD 524.
FT REPEAT 21090 21118 WD 525.
FT REPEAT 21130 21158 WD 526.
FT REPEAT 21170 21198 WD 527.
FT REPEAT 21210 21238 WD 528.
FT REPEAT 21250 21278 WD 529.
FT REPEAT 21290 21318 WD 530.
FT REPEAT 21330 21358 WD 531.
FT REPEAT 21370 21398 WD 532.
FT REPEAT 21410 21438 WD 533.
FT REPEAT 21450 21478 WD 534.
FT REPEAT 21490 21518 WD 535.
FT REPEAT 21530 21558 WD 536.
FT REPEAT 21570 21598 WD 537.
FT REPEAT 21610 21638 WD 538.
FT REPEAT 21650 21678 WD 539.
FT REPEAT 21690 21718 WD 540.
FT REPEAT 21730 21758 WD 541.
FT REPEAT 21770 21798 WD 542.
FT REPEAT 21810 21838 WD 543.
FT REPEAT 21850 21878 WD 544.
FT REPEAT 21890 21918 WD 545.
FT REPEAT 21930 21958 WD 546.
FT REPEAT 21970 21998 WD 547.
FT REPEAT 22010 22038 WD 548.
FT REPEAT 22050 22078 WD 549.
FT REPEAT 22090 22118 WD 550.
FT REPEAT 22130 22158 WD 551.
FT REPEAT 22170 22198 WD 552.
FT REPEAT 22210 22238 WD 553.
FT REPEAT 22250 22278 WD 554.
FT REPEAT 22290 22318 WD 555.
FT REPEAT 22330 22358 WD 556.
FT REPEAT 22370 22398 WD 557.
FT REPEAT 22410 22438 WD 558.
FT REPEAT 22450 22478 WD 559.
FT REPEAT 22490 22518 WD 560.
FT REPEAT 22530 22558 WD 561.
FT REPEAT 22570 22598 WD 562.
FT REPEAT 22610 22638 WD 563.
FT REPEAT 22650 22678 WD 564.
FT REPEAT 22690 22718 WD 565.
FT REPEAT 22730 22758 WD 566.
FT REPEAT 22770 22798 WD 567.
FT REPEAT 22810 22838 WD 568.
FT REPEAT 22850 22878 WD 569.
FT REPEAT 22890 22918 WD 570.
FT REPEAT 22930 22958 WD 571.
FT REPEAT 22970 22998 WD 572.
FT REPEAT 23010 23038 WD 573.
FT REPEAT 23050 23078 WD 574.
FT REPEAT 23090 23118 WD 575.
FT REPEAT 23130 23158 WD 576.
FT REPEAT 23170 23198 WD 577.
FT REPEAT 23210 23238 WD 578.
FT REPEAT 23250 23278 WD 579.
FT REPEAT 23290 23318 WD 580.
FT REPEAT 23330 23358 WD 581.
FT REPEAT 23370 23398 WD 582.
FT REPEAT 23410 23438 WD 583.
FT REPEAT 23450 23478 WD 584.
FT REPEAT 23490 23518 WD 58
```



```

RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Katayama S., Kitamura K., Toda T.;
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Baddock K., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AB032410; BAA84528.1; -
CC EMBL: Z94864; CAB08168.1; -
CC InterPro: IPR001680; -
CC InterPro: IPR001810; -
CC Pfam: PF00646; F-box.1.
CC Pfam: PF00400; WD40; 7.
CC PRINTS: PR00320; GPROTEINBPT.
CC PROSITE: PS50181; FBOX; 1.
CC PROSITE: PS00678; WD_REPEATS_1; 2.
CC PROSITE: PS50082; WD_REPEATS_2; 7.
CC PROSITE: PS50294; WD_REPEATS_REGION; 1.
CC Repeat: WD repeat.
CC KW DOMAIN 107 153 F-BOX.
CC FT REPEAT 271 299 WD 1.
CC FT REPEAT 311 339 WD 2.
CC FT REPEAT 350 379 WD 3.
CC FT REPEAT 390 420 WD 4.
CC FT REPEAT 432 460 WD 5.
CC FT REPEAT 472 500 WD 6.
CC FT REPEAT 510 538 WD 7.
CC SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match 19.5%; Score 590.5; DB 1; Length 605;
Best Local Similarity 29.0%; Pred. No. 1.9e-36;
Matches 156; Conservative 86; Mismatches 189; Indels 107; Gaps 12;

OY 80 SSATVKKRKLSAYKEKELCVKTFEOMSE---SDQVEFEHLISOMCHYGHINSTL 136
DB 46 SSM---HNELSGLSKSRQREAVAAAFSEASCSEKRLALOGILNCCSSLSLFASTL 101
OY 137 KPMLODFITALPARGLDHIAENILSYLDAKSLCAAEIWCKEVRYTSDGMLKRLIERH 196
DB 102 DSLVRLDPSLFLV---EISFRILSLFKAKSLQAAOYSKHKEIADDDVLIHRCBOH 157
OY 197 VRTDSLMRGLAERNGGOYLFXN-----
DB 158 INNK-----CEKGMGLPLENLTLYAAKASTOKRYERLTRKGVQAOAHSSPVKAKLD 211
OY 220 -----KPPDGNAPNSF-----YRALYPKTIIDTETIESNMRCGRHSIQ 258
DB 212 DYEPTSSNEETISSVKKPPSPNSDSKFLPFRTRPWKEVYERCR---VECNMRHGR--- 263
OY 259 RHCR-----SETSKGYVCLQYDQKIVSGRNDTIKIMDKNTIECKRIILHGSVYCLQ 314
DB 264 ---CROYVLSGHDGVMCLQVLRNIIASGSYDTIRLMNLATQOVALLEGHSVGTCLQ 320
OY 315 YDERVITIGSSDSTVRVMDVNTGEMLTILHCEAVLHLFFNNGMVATCSKDRSIAMWM 374
DB 321 FDCKILISGMDKTIIRIMNRTSECSILHGHDSVLCLEFDSTLLVSGSADCTVAKMHF 380
OY 375 ASPTDITLRVVLGHRAAVVDF---DDKTIYASGDRTIKVNTSTCEVFRLLNGHKRG 432

```

```

DB 381 SGGRITLR-----GHTGPVNSVRIIPDRGLVLSGSDSTIKTMSLETNCLHFFSAHGP 436
OY 433 IACLOYRDLVYSGSSDNTIRLMDIECAGCLRYLECHEEYLRRCIRPNDRKIVSGAYDKI 492
DB 437 YOSLALADRLPSCSLDGTIRKQWIDIEKKCKVHTLFFHIGVEIADHLRLISGADHGVY 496
OY 493 KWDVLAALDPAPRACTLCRLTVEHSGVRFRLQFDFEPIVSSHDDTLINDELNDP 550
DB 497 KWEEACE-----CVHTLKNHSEPTVSVALDCVSGSEDEKITYLMLFNAP 543

RESULT 6
SCOB_EMENT STANDARD; PRT; 678 AA.
ID SCOB_EMENT
AC 000659;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DE 01-OCT-2000 (rel. 40, Last annotation update)
DE SULFUR METABOLITE REPRESSION CONTROL PROTEIN.
GN SCOB OR MAPB1.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiates; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RA Natorff R.;
RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE
CC REPRESSION.
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MET30/SCOB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U21220; AAC15905.1; -
CC InterPro: IPR001680; -
CC InterPro: IPR001810; -
CC Pfam: PF00646; F-box.1.
CC Pfam: PF00400; WD40; 6.
CC PRINTS: PR00320; GPROTEINBPT.
CC PROSITE: PS50181; FBOX; 1.
CC PROSITE: PS00678; WD_REPEATS_1; 4.
CC PROSITE: PS50082; WD_REPEATS_2; 7.
CC PROSITE: PS50294; WD_REPEATS_REGION; 1.
CC Transcription regulation; Repeat; WD repeat.
CC KW DOMAIN 178 224 F-BOX.
CC FT REPEAT 347 375 WD 1.
CC FT REPEAT 387 415 WD 2.
CC FT REPEAT 427 455 WD 3.
CC FT REPEAT 466 496 WD 4.
CC FT REPEAT 508 534 WD 5.
CC FT REPEAT 553 595 WD 6.
CC FT REPEAT 607 635 WD 7.
CC FT REPEAT 647 675 WD 8.
CC SEQUENCE 678 AA; 76070 MW; D840DA452E37BAC53 CRC64;

Query Match 19.0%; Score 575; DB 1; Length 678;
Best Local Similarity 26.9%; Pred. No. 3.1e-35;
Matches 149; Conservative 83; Mismatches 168; Indels 154; Gaps 14;

OY 117 VEHLISOMCHYGHINSTYKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEIWC 176
DB 153 LOGIMACCPOLSYISATVRLRIDFTALP-----PETAIRILCYLDITTSICKASOVS 208

```

```

QY 177 KEMVRYTSDGMLKKLIERAVRTDSLRLGLAERGGQYLFKNK----- 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 RGRRLADDDVYVHNRCEQIHRK-----CKKCGMLPLDRKRLRESKRELELATW 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 -----PPDGA--PP-----NSFYRALY---P 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 263 DKGVGPRSDASPSGSKRLKLEDEVAVVRHCSLSGSDAGVDDSDFFRYRPMK 322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 238 KIIODIETIESNMRCGRHSIORHCRSESK-----GVYCLQYDDOKIYSLRNTIKIMD 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 EVYKDRKVTNMYKGR-----CSTKTFKHTNGVMCLQFEDNLILAGSYDTTIKIMD 375
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 KNTLECKRLITGHTGSLVCLQYDERVLIITSSSTFVRVMDVNGEMLNTLIHCEAVLH 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 TEGEGLRLTGRHESGRICQFDPDKLISSMRTIKVMMWRKGECLISTYTGIRGVIGL 435
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 RENNMMVNTCSKDRSLAVMDASPTDITLRLVLGHRRAVNV--VDEDKYIVASGDRT 411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 436 HFDASILASGSDVKTYKIVNFEDEKSTFSLR-----GHTDWNVAVRDTSSKTVFSASDDCT 491
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 412 IKVMNNTSGCFVRLTNGH-----KRGIAAC----- 435
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 492 VRLMDLDTKTCITRTEHGHGQVQVPLPREFEFEHDAECENDLSTSGDANPSSIOA 551
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 436 -----LQYRD-----RLVYSGSSDNTIRLMDIEGACLRVLEGHEEL 472
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 552 SMGLEPRAVASQSASFSTFNDNGRAAPRRYMTSALDSTRIMETTTGRLRFFGHLS 611
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 473 VRCIRPDNRKIYSGAIDGKTKVMDLVAAIDPRAPACTLCLRTLVESGRVRLQDEFOI 532
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 612 VMALGADTLRLTIVGSAEDRMKIMD-----PRTGKCEPFTGSHGVPYTCIGLDSRF 662
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 533 VSSSHDITILIMDF 546
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 663 ATGSEDCVARMYSF 676
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE MET30/SCONE/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 246861; CAA86905.1; -.
DR EMBL: 126505; AAA96717.1; -.
DR SGD: S0001308; MET30.
DR InterPro: IPR001800; -.
DR InterPro: IPR001810; -.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 6.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PROSITE: PS50181; PROX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 4.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR KX Transcription regulation; Methionine biosynthesis;
DR Cysteine biosynthesis; Repeat; WD repeat.
FT DOMAIN 181 227
FT REPEAT 300 328 WD 1.
FT REPEAT 340 368 WD 2.
FT REPEAT 380 408 WD 3.
FT REPEAT 419 449 WD 4.
FT REPEAT 461 499 WD 5.
FT REPEAT 509 538 WD 6.
FT REPEAT 550 578 WD 7.
FT REPEAT 607 635 WD 8 (POTENTIAL).
FT CONFLICT 61 61 M -> I (IN REF. 1).
SQ SEQUENCE 640 AA; 72835 MM; 5135D4BCA2E1EB97 CRC64;

Query Match 18.0%; Score 545; DB 1; Length 640;
Best Local Similarity 28.9%; Pred. No. 4.8e-33;
Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

QY 34 IPEKSLRQTYNSCARLILN-OETVCLASTAMKTENCYAK-TKLANGTSMIVKORLKS 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 86 LPEVNFPTKCYVHNNDIDFSPRHTACYKODLKRQEIINANAKLPLQDSOIHNIISKYS 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 92 ASYEKEKELCYAFPEOMSESDOVERFENHILISOMCHYOGHINSLYKRLQDFITALPAR 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 146 NSNDKIRKL-----ILDGILSTSCPPOLSYISLVTNMIKIDFISILP-- 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 GLDHAENILSYLDAKSICAAELVCKEYRVATSDGMLKKLIERAVRTDSLRLGLAERR- 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 --QELSLKILSTLSDCOSLCNATRVCKRMOKLADDDRVYHNCEDHIT-----DKKC 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 211 ---GNGOYLFKNK-----PPDGNAPNSFYRALYPKTIODIETIESNMRCGRHSIQ 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 RYCGMGGLPLHMKRRARIQONSTGSSNADIDQTRPRPKYIVREFKYESMMRGG----- 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 259 RIHGRSETSK-----GVYCLQYDDOKIYSGLDNDNTIKIMDKNTLECKRLITGHTSVCLO 314
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 292 --HCRIOEFKGMHGDVLTLQFENYRLFLFGSYDSTIGIDLFTGKILRLISGHSQGVKTY 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 315 YVERIITIGSSSYRVAVMDVNGEMLNTLIHCEAVLHRLRNNNGMMVNTCSKDRSLAVMD 374
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 FDDRLLITIGSLDKTRVNNYITGECISTYRGHSVLSVDSYQAVIYSGSDAKTVKWHV 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 375 ASPTITLRLRVLGHRAVNVVDFDKYI--VSASGDITKVMWTSCEAFVRLTNGH--- 429
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 410 ESRITYTLR-----GITEVNVCKYKILHFKSFSCSCSDDTITMMQDIRNSCLKYVRGHGQ 465
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 430 KRGIACLQYRD--RLVYSGSS----- 448

```



```

Db 466 VQITIPLTIKDVENLATDNTSDGSSPODDPTMTDGADESPTSPNEQETVLNIDENIPYPTHL 525
Qy 449 -----DNTIRLMDIEGACALRYLEGEHELRYCIRFNKRIVSGAYDGKIKVMDLVAALDP 503
Db 526 LSCGLNITIKIMOVKIKGKCRIRFOGHVEGWMDIAADNFRITSSHGSGIKVMDLQSG---- 582
Qy 504 RAPAGTLCRLTVEHSGRVFRLOFDEFQIVSSSHDDTI 541
Db 583 -----KCMHTE---NGR--RLQRETOHTOTOSLGDKV 609

RESULT 8
SC02_NEUCR STANDARD; PRT; 650 AA.
AC 001277;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE SULFUR CONTROLLER-2 (SCON2).
GN SCON-2.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A;
RX MEDLINE=95241499; PubMed=7724564;
RA Kumar A., Paletta J.V.;
RT "The sulfur controller-2 negative regulatory gene of Neurospora
RT crassa encodes a protein with beta-transducin repeats.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).
CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
CC -1- INDICATION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MET30/SCOMB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U17251; AAA68968.1; -
CC InterPro: IPR001680; -
CC DR InterPro: IPR001680; -
CC DR Pfam: PF00646; F-box; 1.
CC DR Pfam: PF00400; WD40; 7.
CC DR PRINTS: PR00320; GPROTEINBRPT.
CC DR PROSITE: PS50181; FBOX; 1.
CC DR PROSITE: PS50678; WD_REPEATS_1; 2.
CC DR PROSITE: PS50082; WD_REPEATS_2; 6.
CC DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
CC KW Transcription regulation; Repeat; WD repeat.
FT DOMAIN 124 170 F-BOX.
FT REPEAT 292 320 WD 1.
FT REPEAT 332 360 WD 2.
FT REPEAT 372 400 WD 3.
FT REPEAT 411 441 WD 4.
FT REPEAT 453 488 WD 5.
FT REPEAT 528 564 WD 6.
FT REPEAT 576 604 WD 7.
FT REPEAT 616 644 WD 8.
SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

```

Query Match 17.5%; Score 531.5; DB 1; Length 650;
Best Local Similarity 25.0%; Pred. No. 4.9e-32;

```

Matches 144; Conservative 76; Mismatches 180; Indels 177; Gaps 11;
Qy 117 VEHLISOMCHYOHGHINSYLKPMLOPDTALPARGLDHAENILSYLDKSLCAALVC 176
Db 99 LOGILSOLCEPQLSPFVSREYNELKIDFISALPV-----ELAQVLYCLDPTVSLTKAQS 154
Qy 177 KEWRYVTSOGMLKKLIERVYRTDSLMRGLAERGGQYLFKKP-----P 222
Db 155 QWRRTLADSDAVVWRCOEQVNRK-----CTKCGGLPLERKKLRNYYTOROLAKGG 208
Qy 223 DGNA-----PNSFYRA 224
Db 209 QGRVTELADSHDSQDSVNOHGRKPAEAEEEDPIKKRQCMAAEASKAVTQPKTSMKA 268
Qy 235 LYPKIIODIETTESNMRGSHSLQRIHCRSEISGKYCYQYDQKIVSGLRDNTIKIMDK 294
Db 269 VY-----RDRQVSYNNKNSKYLSTL---KGHNQVTCQLQDNLNLTATSGYDTITIMNI 321
Qy 295 NTLCKRILTGHTGSLVLCQYDERVYITIGSSDSTVWMDVNTGEMLNTLIHCEAVLHLR 354
Db 322 ETEECIRTLVGHGTAGIRALQFDQSKLISGSLDHTIKVMNHTGECSTFRAHDSVSYVA 381
Qy 355 FNNGMVWTCSDKRSIAVMMASTDTLRVLYGHRAAVNV--VDDDKYIVASGDRIT 412
Db 382 FDGHLIASSGSSDKTVKIFDPSKEYCYCLK---GHSQWNVSTHVDIKSRVFSASDPTI 437
Qy 413 KYVNTSTCEFEVFTLNGHKRCI-----ACIQYDRILVVS-----445
Db 438 KLMDLDTROYIRYIEGHVGHVQVLLIPPEYEDDEVYLNASQDNQDAMSVSGSGSPS 497
Qy 446 -----GSSD 449
Db 498 MSHAQIERAGSPGSHSSHNLPSLPSCGDEVRHLYGSAFVADESRLPPRYFMGGLD 557
Qy 450 NTRIRLMDIEGACALRYLEGEHELRYCIRFNKRIVSGAYDGKIKVMDLVAALDPAPAGI 509
Db 558 STWRLDSATGRLRLTFLGHEGWSLAGPTIRIVISGANQKWE-----PRSG 608
Qy 510 LCURTVEHSGRVFRLOFDEFQIVSSSHDDTIIMPF 546
Db 609 KCDATYTGHCGRPYTCGLSDSLMASGSEDTIRLHSF 645

RESULT 9
SE10_CAEEL STANDARD; PRT; 579 AA.
AC 093784;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE SEL-10 PROTEIN.
GN SEL-10 OR F55B12.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Sims M.
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RL [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA Jones S.J.M.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```



```

DT 01-OCT-2000 (rel. 40, last annotation update)
DE CELL DIVISION CONTROL PROTEIN 4.
GN CDC4.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
RN NCBI_TaxID=5476;
  [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SGY126;
RA Shieh J.C., White A.M., Rosamond J.;
  Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
CC VARIOUS ASPECTS OF SPOREATION. REQUIRED FOR HET1-HTB1 LOCUS
CC TRANSCRIPTION ACTIVATION (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X96763; CAA65538.1; -
DR InterPro: IPR001680; -
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 4.
DR PROSITE: PS50082; WD_REPEATS_2; 4.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
FT DOMAIN 212 258 F-BOX.
FT REPEAT 322 351 WD 1.
FT REPEAT 363 391 WD 2.
FT REPEAT 403 431 WD 3.
FT REPEAT 442 468 WD 4 (POTENTIAL).
FT REPEAT 478 506 WD 5.
FT REPEAT 519 549 WD 6.
FT REPEAT 561 589 WD 7.
SQ SEQUENCE 684 AA; 76090 MW; 3DD65DB31293B107 CRC64;

```

Query Match 15.7%; Score 477.5; DB 1; Length 684;
 Best local similarity .25.4%; Pred. No. 5,3e-28;
 Matches 136; Conservative 95; Mismatches 184; Indels 121; Gaps 16;

```

DB 399 HIFRGHTSTIRCLDIHPAVIGNKQDGEDIVPEEYLLITGSRDHNHAWKLPVYDSDS 458
OY 375 -----ASPTDITLR-----RVLY 387
DB 459 YIETEGEEDNPPLIYVLSGHQSVASISGYNIITSSYDSIVRWDLDDGCHTHVQ 518
OY 388 GH--RAAVNVVDEDDKYIVSASGRTIKYWNSTCEFEVTLNGKRGKIACIYRDLVVS 445
DB 519 GHLDVRYSTALDPHSKTCPSGMSDSNIWNVNFETGLKKVLYGHASLVGLDLYDDVVS 578
OY 446 GSSDNIIRIMDIECGACIYVLSGHELVLCIRPDNKRIVSGAYDKIKYWDVLAALDPA 505
DB 579 AADATLRIMDKATGLKRGKLGHAITCFEHDGLRVVSGS--EKMLKLMV----- 629
OY 506 PGTLICFTLVESGVRFLQD-----EPOIVSSHDDTLT-TWDFLNDPAAQ 553
DB 630 EKGTFARDLSDVTGSGIMQYREDYKRCVAAVQRITNEDEGETTIELDF-SQPLIQ 684

RESULT 12
HET1_PODAN STANDARD; PRT: 1356 AA.
AC Q00808;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE VEGETABLE INCOMPATIBILITY PROTEIN HET-E-1.
GN HET-E1.
OS Podospota anserina.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Podospota.
OX NCBI_TaxID=5145;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96009891; PubMed=7557402;
RA Saube S., Turc B., Begueret J.;
RT "A gene responsible for vegetative incompatibility in the fungus
RT Podospota anserina encodes a protein with a GTP-binding motif and G
RT beta homologous domain."
RL Gene 162:135-139(1995).
CC -1- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH
CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,
CC HET-C.
CC -1- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L28125; AAA85775.1; -
DR InterPro: IPR001680; -
DR Pfam: PF00400; WD40; 10.
DR PRINTS: PR00320; GPROTEINRPT.
DR PROSITE: PS00678; WD_REPEATS_1; 10.
DR PROSITE: PS50082; WD_REPEATS_2; 10.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW GTP-binding; Repeat; WD repeat.
FT NP_BIND 300 307 GTP (POTENTIAL).
FT REPEAT 839 869 WD 1.
FT REPEAT 881 911 WD 2.
FT REPEAT 923 953 WD 3.
FT REPEAT 965 995 WD 4.
FT REPEAT 1007 1037 WD 5.
FT REPEAT 1049 1079 WD 6.
FT REPEAT 1091 1121 WD 7.
FT REPEAT 1133 1163 WD 8.
FT REPEAT 1175 1205 WD 9.
FT REPEAT 1217 1247 WD 10.
SQ SEQUENCE 1356 AA; 149765 MW; 965FB319844E0651 CRC64;

```

Query Match 15.0%; Score 455.5; DB 1; Length 1356;
 Best Local Similarity 26.0%; Pred. No. 5,5e-26;
 Matches 164; Conservative 102; Mismatches 220; Indels 145; Gaps 30;

QY 28 EPPRKIIPEKNSLRQYNSCAR-LCLNQEYVCLASTAMK-----TENCVAKTKLANGTSS 81
 DB 581 EMPDVSDEDSLEIEIKGSGFLIERIRYEVHQSAKDFLGTASDKSNKASQCAFE 640
 QY 82 MIYK-----QRLASYTEKE--LCYKFEQMSQSEDOVEVEHL--ISQCHY 127
 DB 641 LVFPTGEDVSYIIFFWSLWMSOKLRDYLCLNAPGLIDNVRPPDLATVRYSCY 700
 QY 128 QHGINSYLRP-----MLORD-----FIT-----ALPARGLDHI--AEN 159
 DB 701 WIDLRLRLVSTSSKMYHLLQDDGDHRLFTTKYLYLWLEALSLLRALP-EGINAIROLES 759
 QY 160 ILSTLDAKSLCAELVCKEYRVTSDGMLKKLIER--MYRDSLWRGLAERRGQYL 216
 DB 760 LIGHTRIGRLIA--IVRDGYR--ALSYRMITEKAPLQATYSAL-----V 800
 QY 217 FKNKPPGNNAPNSFYRALY--PKIIQDIETIESMWRGRHSL----- 257
 DB 801 F-----APTDSMIKKIRKEEPGNIISTISVEAEAMNACTOTLEHGSSVLSVAFSAD 852
 QY 258 -QRIHCRS-----ETSKG-----VYCLQY--DDOKIYSGLRDNTIKIMDK 294
 DB 853 GQRASSGDDKTIKIMPTASTGTOTLEHGSGSVMSVAFSPDERKVASGDDTKIKIMDA 912
 QY 295 NTLCKRILFGHNGSVLCLOY--DERVIITGSSDSIVRVMDVTGEMLNTLIHHCVAVLH 352
 DB 913 ASGTCTQTLLEGHGGRVQSVAFSPDQGRVASSDDHTIKIMDASGTCTQTLLEGHGSVLS 972
 QY 353 LREN-NGMAYTC-SKDSIAVWMDASPTDITLRVLYVGHAAVAVDF--DQKIYVSA 408
 DB 973 VAFSPDQGRVASSGDDTKIKIMPTASG--TCTQTLLEGHGSVMSVAFSPDQGRVASSD 1029
 QY 409 DRTIKVWNTSTCEVFRTLNHGKRGIACTOYR--DRLVVGSSDPTITLMDIEGACLRVL 466
 DB 1030 DKTIKIMPTASTGTOTLEHGSGVQSVSPDQGRVASSDDHTIKIMVAVSGTCTQTL 1089
 QY 467 EGHLELVKICRF--DNKRIYSGAYDGRIKIMDLVLAALDPRAPAGTCLRLTVEHSGVFR 524
 DB 1090 EGHDSVMSVAFSPDQGRVASSGIDGTIKIMD-----AASGT-CTQTLLEGHGVHVS 1140
 QY 525 LQF--DEFOYVSSHDFTIILMPFLNDPAAQ 553
 DB 1141 VAFSPDQGRVASSGIDGTIKIMDASGTCTQ 1171

RESULT 13
 POP1_SCHPO STANDARD; PRT; 775 AA.
 AC P87060;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE WD-REPEAT PROTEIN POP1.
 OS POP1 OR SPBC262.18.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID:4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97347242; PubMed=9203581;
 RA Komlami K., Toda T.;
 RT "Fission yeast WD-repeat protein pop1 regulates genome ploidy through
 ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rum1
 and the S-phase initiator Cdc18.";

RL Genes Dev. 11:1548-1560(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wood V., Rajadream M.A., Barrell B.G., Devlin K., Churcher C.M.;
 RL Submitted (Mar-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN MAINTENANCE OF PLOIDY THROUGH PROTEASOME
 CC DEPENDENT DEGRADATION OF CDK INHIBITOR RUM1 AND S-PHASE INITIATOR
 CC CDC18. FUNCTIONS AS A RECOGNITION FACTOR FOR RUM1 AND CDC18, WHICH
 CC ARE SUBSEQUENTLY UBIQUITINATED AND TARGETED TO THE 26S PROTEASOME
 CC FOR DEGRADATION.
 CC -1- SUBUNIT: BINDS TO CDC18.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC
 CC EMBL: Y08391; CAA69671.1; -
 CC EMBL: AL022103; CAA17898.1; -
 CC InterPro: IPR001680; -
 CC InterPro: IPR001810; -
 CC Pfam: PF00646; F-box; 1.
 CC Pfam: PF00400; WD40; 6.
 CC PRINTS: PR00320; GPROTEINBPT.
 CC PROSITE: PS50181; FBOX; 1.
 CC PROSITE: PS00678; WD_REPEATS_1; 3.
 CC PROSITE: PS50082; WD_REPEATS_2; 4.
 CC PROSITE: PS50294; WD_REPEATS_REGION; 1.
 CC Repeat; WD repeat.
 CC FT DOMAIN 298 345 F-BOX.
 CC FT REPEAT 444 472 WD 1.
 CC FT REPEAT 484 538 WD 2.
 CC FT REPEAT 575 603 WD 3.
 CC FT REPEAT 615 645 WD 4.
 CC FT REPEAT 657 687 WD 5.
 CC SQ SEQUENCE 775 AA; 87816 MW; B06EBA6553BE01 CRC64;

Query Match 14.9%; Score 453; DB 1; Length 775;
 Best Local Similarity 25.8%; Pred. No. 4,1e-26;
 Matches 162; Conservative 86; Mismatches 229; Indels 150; Gaps 20;

QY 9 QEKALKFNSSSEREDC-----NNGEPPRKIIPEKNSLRQY-NSCARLCLNQE 56
 DB 121 RECKLRKRRSSLSNLHANKRFLFNSQSDGNKKNETFPSTNSVNFYNNCDSEVASET 180
 QY 57 VC-----LASTAMKTENCVAKTKLANGTSMIYPRQRKLSA----- 92
 DB 181 TFSLDAPNNSVNYSYFSPNLLGNDKTKROSFPFSSSSSHNSLHPVLIYDPSSENPISHP 240
 QY 93 -----SYEKKEKL-----CVKYFEQMSQSEDOVEVEHLISQCHYQHGINSYLRKPLQDF 144
 DB 241 SNHLSQKNAVLKLAQLISSFEKLPESVRQYLLFHLILSRGCHHANYONHKLILPFGQNF 300
 QY 145 ITALPARGLDIAENILSYLDAKSLCAELVCKEYR-VTSDGMLKKL-----I 193
 DB 301 LIGFPA-----FITNLVLTFLDAPSLCAVSAQVSHHWYKLVSSNEELMKSILFLKDGFFMDSI 356
 QY 194 EKNVPTDSLWRGLAERRGQYLFKNKPPDGNAPNSFYRALYPRKIIDIETIESMWRG 253
 DB 357 DSKIRTMCEQSLA-----CATMKRVYRHHFLNRRW--- 389
 QY 254 RSLQRIHCRSSTSKGVYC---LOYDQKIYSGLRDNTIKIMDKNTLCKRILFGHNGSV 310
 DB 390 LIAPEIKRCSFPHGVRLITKLOFDDDKLIIVTSCSPRINITYDTFTVGLISLSEHEDV 449
 QY 311 LCLQYDERVIITGSSDSIVRVMDVTGEMLNTLIHHCVAVLHLENNGMAYTCSDKNSIA 370

```

DB 450 WFEFVGDITLVGSDRTVWDLRTGECKOYFEGTSTIRCKIYQGNQSTDTD----- 505
OY 371 VMDMSPDTILRLVGRVAVNVDPDDKIVASGSRITKVMN----- 416
DB 506 -----DVEKE-----NRPASNDANSMPYITISSRCDTIRLWLPCLDDPPVYVNE 552
OY 417 -----TSTCEFYVTLNGHK-----RGIACTLOYRDLVYSGSSDNTIRLMDIECGAC 462
DB 553 NPDQNDNFTSATINPEYITRLGHNDVSREVACL---GLIIVASAYDGLRLVWKASTGYC 609
OY 463 LRVLSGHELVLCIRFDNKR--IVSGAYDGKTKVMDLVALDPRAPAGTLCRLTVEHSG 520
DB 610 LHVLRGHWVRVYSVTINPFRQOCISAGTDKIRIMWL-----ESGEL-LQTLIGHSN 660
OY 521 RVFRLOFDEFQIVSS--HDDTILMD 545
DB 661 LVSVTFNONIIVSASAPDTSILRWMD 687

RESULT 14
CC4_YEAST STANDARD: PRT: 779 AA.
AC P07834;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CEL DIVISION CONTROL PROTEIN 4.
GN CDC4 OR YFL009W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86011240; PubMed=3309335;
RA Yocham J., Byers B.;
RT "Structural comparison of the yeast cell division cycle gene CDC4 and
RT a related pseudogene."
RL J. Mol. Biol. 195;233-245(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murkamt Y., Naitou M., Hagihara H., Shibata T., Ozawa M.,
RA Sasunuma S.-I., Sasunuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae."
RT Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 1-579 FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrell B.G., Churcher C., Rajandream M.A.;
RT Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
RL -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
RL REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
RL POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
RL DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
RL VARIOUS ASPECTS OF SPOULATION. REQUIRED FOR HTAI-HTBI LOCUS
RL TRANSCRIPTION ACTIVATION.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC *-----
CC EMBL; X05625; CAA29113.1;

```

```

DR EMBL; D50617; BAA09229.1; -
DR EMBL; D31600; BAA06495.1; -
DR EMBL; 246255; CAA86341.1; -
DR PIR; A26867; A26867.
DR SGD; S0001865; CDC4.
DR InterPro; IPR001680; -.
DR InterPro; IPR001810; -.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 6.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
FT DOMAIN 272 319 F-BOX.
FT REPEAT 380 408 WD 1.
FT REPEAT 420 449 WD 2.
FT REPEAT 461 493 WD 3.
FT REPEAT 528 556 WD 4.
FT REPEAT 568 598 WD 5.
FT REPEAT 630 658 WD 6.
FT REPEAT 669 698 WD 7.
FT CONFLICT 460 460 K -> E (IN REF. 1).
SQ SEQUENCE 779 AA; 86089 MW; 0348F2F8FA78F3BC CRC64;

```

Query Match 13.2%; Score 399; DB 1; Length 779;

Best Local Similarity 24.3%; Pred. No. 4,1e-22; Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

```

OY 56 TVCLASTAMKTCNVAKTRLANGTSSMIVPKOKRISASYEKER-----ELCVKFEQW 108
DB 194 TPLPLAKTKTINN-----NNNIADLESKDSIISEYLSDFEISANNMLPHAYFK-- 244
OY 109 SESDQVEFEVHLISQCHQHGHINSYLRPMLOQDITPLPARGLDIAENLISYDAAS 168
DB 245 -----NLFRVLVANNDRSELSDGLTLDNLRDITSLP-----EISLKTFFNYLQFED 294
OY 169 LCAELVCKEWY-VPSDMLMKRLI--BRMYRTDSLMLGLAERRGQGYLFKNKPPDGN 225
DB 295 IINSLGVSONWKKIIRKISLMKKLISENFV-----SPKGF 331
OY 226 APPNSFYRALYPTIID-----IET--TESNWRGGRSLORICRSETSKGYCLQYD 276
DB 332 NSLNLKLSQKRYPLSQODRLSLFLENIFLKNWYNPKFVPOFTLRGHMTSVITCLOPE 391
OY 277 DQKIVSGLDNFKIKDKNTLECKRILTLGHTGSLQLODE-RVITGSSDSIVRVWDVN 335
DB 392 DNYVITGADDKMIRVYDSINKKFLQLQSGHDGVMALKYAHGGLIVSGSDRTVRWMDIK 451
OY 336 TGEMLNTLIHCEAVLHLRFNNQNMVTCCKDRSIAVMDASPTDILRLVGRVAVNV 395
DB 452 KG-----CCF-----HYFGHNSIVVC 468
OY 396 VFEDD--KYIVSASGDRITKVMN-----STCE---FVETLNG 428
DB 469 LDIVERKNIKVIYTGSRDNTLHWKLPKSSVDPDHEEDHYPLVFTPRENPFFVGVCLNG 528
OY 429 HRRGIACLOYRDLVYSGSSDNTIRLMDIECGACLVLSGHELVLCIRFDN--KRIYVG 486
DB 529 HNASVTVSGHGNIIVSGSDNTLIWMDVAOMKCLYILSGHDIRIYSTIYDHRKRCISA 588
OY 487 AVDGKIKVMDL-----VALDPRAPAGT--CLRTLVHSGVFRLOFDEFQIVSS 535
DB 589 SMQTTIRIMDLNINWNGECSTATNSASPCAKILAGMTLLOGHTALVGLRLSDKFLVSA 648
OY 536 SHDDTILMDFLNDPAAQAEPPSPSRITYY 566
DB 649 AADGSIRGMD-AND-----YSRKFSY 668

```

RESULT 15

POP2_SCHPO STANDARD: PRT; 703 AA.

ID POP2_SCHPO 014170;

AC 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE WP-REPEAT PROTEIN POP2 (PROTEOLYSIS FACTOR SUD1).

GN POP2 OR SUD1 OR SPAC4D7.03.

OS Schizosaccharomyces pombe (Fission Yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

XX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA Wolf D.A., Jackson P.K.;

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=98318628; PubMed=9653157;

RA Jallepalli P.V., Tien D., Kelly T.J.;

RT "sudd" targets cyclin-dependent kinase-phosphorylated Cdc18 and Rrm1

RT proteins for degradation and stops unwanted diploidization in fission

RT yeast.";

RL Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA Gentiles S., Churcher C.M., Bartell B.G., Rajandream M.A., Wood V.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: PUTATIVE COMPONENT OF E3 (PROTEIN-UBIQUITIN LIGASE)

CC COMPLEX: TARGETS CYCLIN-DEPENDENT KINASE-PHOSPHORYLATED CDC18 AND

CC RRM1 PROTEINS FOR DEGRADATION.

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.

CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF038867; AAB95480.1; -

DR EMBL: AF064515; AAC39496.1; -

DR EMBL: Z98602; CAB1275.1; -

DR InterPro: IPR001680; -

DR InterPro: IPR001810; -

DR Pfam: PF00646; F-box; 1.

DR Pfam: PF00400; WD40; 6.

DR PRINTS: PR00320; GPOTEINRPT.

DR PROSITE: PS50181; FBOX; 1.

DR PROSITE: PS00678; WD_REPEATS_1; 3.

DR PROSITE: PS50082; WD_REPEATS_2; 6.

DR PROSITE: PS50294; WD_REPEATS_REGION; 1.

DR Repeat; WD repeat.

KW DOMAIN 236

FT REPEAT 389 F-BOX.

FT REPEAT 429 WD 1.

FT REPEAT 505 WD 2.

FT REPEAT 545 WD 3.

FT REPEAT 587 WD 4.

FT REPEAT 625 WD 5.

FT REPEAT 654 WD 6.

SO SEQUENCE 703 AA; 79605 MW; 0336A0568C152396 CRC64;

Query Match 13.1%; Score 396.5; DB 1; Length 703;

Best Local Similarity 23.3%; Pred. No. 5,6e-22;

Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

OY 14 KFMNSERED--CNGGEPRKI-----IPEKNSLRQTYNSCARLCL-----NOETVCL 59

Db 96 KFMNSIENVSFKLDHSPDPDVSFGDFSISLVPQRNFL-YSHSLPPIKSIISDRNRRIKD 154

OY 60 ASIAMKTENCVAKTAKTIANSTSSMIVPKOKKLSASEK-EKELCYVYFEQWSSDDVEEVE 118

Db 155 NTSISNSNDPFPSPKPV-DTSNTVSPGSKPISDEEDLNLOSIVQTFDLPGLQ-SYAF 211

OY 119 HITSQCHYQHGH-INSYLPKMLQDGFALPARGLDHAENILSYLAKSICAAELVCK 177

Db 212 FQLLRSCNQSRMLNCEPILKDI LSNLP-----SIVGSILNLNDIHSLSRLVSP 267

OY 178 EWMYV-----TSDGMLKKLIERNVRTDSLWRGLAERCGOYLFRNKKPPDGAPP-----N 229

Db 268 TWMRLDVHTS--YWKHMF-----SLFGQINENDW-----KYAMPNLRRPPFLHND 312

OY 230 SFYRALYPKIIDIETIEENMRGHSIOR-----

Db 313 QISDDYFPEIFK-----RHFLNRKRWLPSPSPSHLSPPHVPNEMITSLILH 360

OY 260 -----IHRSEFSK-GVYCLQYDDQKIYSGLRDNTIKIMDKNT 296

Db 361 KDRITTSGGSGTIQIHNAITGVLEARLEBKRGKGVAAKHEHTLVSGSIDKTVRWNTIEK 420

OY 297 LECKRILTGHGSSVLCLOY-----DERVIITGSSDSTVYRWVDVTGEML 340

Db 421 AKCTHIFRGHISIRCLLETILVPSRLIRHGVEIYEPDQPIYVSGSRDHLRWV----- 472

OY 341 NTLIHCEAVLHFRFNGMMVTCCKDORSIAVMDMASPTDITLRVLYGHRVAVNYVEDD 400

Db 473 -----KLPKNTDPPYLPDNTNSIDRWE-KNPYFV--HTLGHDSVRIISGIG 517

OY 401 KYVASAGDRTIKVWNTSYCEFEVRLNGHKGRIACLOYRD--RLVVGSSSDNTIRLMDIE 458

Db 518 DILVSGSYDSIRIMRVSGECLYHLRGHSLRIYSVLVEPENICISGSMKSIWMLDS 577

OY 459 CGACLRVLEGHELYRCIFDNKRIVSGAYDGKIKVMDLYAALDPRAGTLCITLYEH 518

Db 578 TGCRKYVLEGHDAFYLLNVPONRLISGSADSTIRIMD--LNTGKP-----LWVLP SN 628

OY 519 SGRVFLQDFEFOIYSSSHDITLIMD 545

Db 629 SGYISSFVSDHKIIT-SGMDGSVKIMD 654

Search completed: June 7, 2001, 10:38:26

Job time: 145 sec

THIS PAGE BLANK (USPTO)

Thu Jun 7 10:44:15 2001

us-09-415-795-4.rsp

Page 13

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 10:35:02 ; Search time 33.7 Seconds
(without alignments)
1978.970 Million cell updates/sec

Title: US-09-415-795-4

Perfect score: 3034
Sequence: 1 MDPAEAVLQEKALKFNNSS.....PAAQAEPPRSPSRITYISR 569

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP:archaea:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP:invertebrate:*
6: SP:mammal:*
7: SP:mhc:*
8: SP:organelle:*
9: SP:phage:*
10: SP:plant:*
11: SP:rodent:*
12: SP:unclassified:*
13: SP:vertebrate:*
14: SP:virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3034	100.0	569	4 Q9Y213	Q9Y213 homo sapien
2	3006	99.1	605	4 Q9Y297	Q9Y297 homo sapien
3	2997	98.8	569	11 Q9QU15	Q9QU15 mus musculu
4	2990	98.5	569	11 Q9Z159	Q9Z159 mus musculu
5	2978	98.2	569	11 Q9R1G7	Q9R1G7 mus musculu
6	2400	79.1	529	4 Q9P2S8	Q9P2S8 homo sapien
7	2384.5	78.6	542	4 Q9UKB1	Q9UKB1 homo sapien
8	2384.5	78.6	550	4 Q9Y4C6	Q9Y4C6 homo sapien
9	2358	77.7	508	4 Q9P2S9	Q9P2S9 homo sapien
10	2089.5	68.9	510	5 Q44382	Q44382 drosophila
11	2089.5	68.9	510	5 Q9VDE3	Q9VDE3 drosophila
12	690	22.7	506	3 Q9P7V1	Q9P7V1 schizosacch
13	661	21.8	1326	3 Q9VZF4	Q9VZF4 drosophila
14	640	21.1	553	4 Q9NUX6	Q9NUX6 homo sapien
15	520	17.1	587	5 Q44083	Q44083 caenorhabdi
16	365.5	12.0	942	5 Q96611	Q96611 dictyosteli
17	356	11.7	410	13 Q9PTR5	Q9PTR5 gallus galli
18	344	11.3	411	5 Q96698	Q96698 drosophila
19	339.5	11.2	777	3 Q9USN3	Q9USN3 schizosacch

20	334.5	11.0	422	4 Q9UKT8	Q9UKT8 homo sapien
21	334.5	11.0	422	4 Q9UKA3	Q9UKA3 homo sapien
22	332.5	11.0	317	10 Q9M222	Q9M222 arabidopsis
23	332	10.9	454	11 Q9QUH1	Q9QUH1 mus musculu
24	330	10.9	333	10 Q9SY00	Q9SY00 arabidopsis
25	325	10.7	277	11 Q9R2A6	Q9R2A6 mus musculu
26	325	10.7	391	11 Q35392	Q35392 mus musculu
27	324.5	10.7	334	4 Q9NMW7	Q9NMW7 homo sapien
28	324.5	10.7	362	4 Q9UGP9	Q9UGP9 homo sapien
29	313.5	10.3	481	5 Q9VPR4	Q9VPR4 drosophila
30	313.5	10.3	485	4 Q9NVX2	Q9NVX2 homo sapien
31	312.5	10.3	330	4 Q9NUL4	Q9NUL4 homo sapien
32	312	10.3	876	10 Q9LFE2	Q9LFE2 arabidopsis
33	308	10.2	480	5 Q96995	Q96995 drosophila
34	304.5	10.0	1205	4 Q9UJ66	Q9UJ66 homo sapien
35	304.5	10.0	1205	4 Q9UJ59	Q9UJ59 homo sapien
36	301.5	9.9	476	13 Q93531	Q93531 xenopus lae
37	299.5	9.9	1205	4 Q9UJ65	Q9UJ65 homo sapien
38	298.5	9.8	1171	4 Q9UJ58	Q9UJ58 homo sapien
39	298.5	9.8	1205	4 Q9UJ67	Q9UJ67 homo sapien
40	298.5	9.8	1205	4 Q9UJ64	Q9UJ64 homo sapien
41	298.5	9.8	1205	4 Q9UJ63	Q9UJ63 homo sapien
42	297	9.8	502	3 Q74855	Q74855 schizosacch
43	295.5	9.7	1205	4 Q9UJ60	Q9UJ60 homo sapien
44	293.5	9.7	1205	4 Q9UJ62	Q9UJ62 homo sapien
45	293	9.7	760	10 Q9LHN3	Q9LHN3 arabidopsis

ALIGNMENTS

RESULT	ID	Q9Y213	PRELIMINARY:	PRT:	569 AA.
DT	01-NOV-1999	(TREMBLrel. 12, Created)			
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)			
DE	01-JUN-2000	(TREMBLrel. 14, Last annotation update)			
DE	BETA-TRANSDUCIN REPEATS CONTAINING PROTEIN.				
CN	BETA-TRCP OR FBW1A.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LYMPHOID;				
RX	MEDLINE=98325370; PubMed=9660940;				
RA	Margolin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,				
RA	Thomas D., Strebel K., Benarous R.;				
RT	"A novel human WD protein, h-beta Trcp, that interacts with HIV-1 Vpu				
RT	connects CD4 to the ER degradation pathway through an F-box motif.";				
RL	Mol. Cell 1:565-574(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Cenciarelli C., Chlaud D.S., Guardavaccaro D., Parks W., Vidal M.,				
RA	Pagano M.;				
RT	"Identification of a family of human F-box proteins.";				
RL	Curr. Biol. 9:1177-1179(1999).				
DR	EMBL: Y14153; CAA74572.1; -.				
DR	EMBL: AF129530; AAF04464.1; -.				
DR	INTERPRO: IPR001800; -.				
DR	INTERPRO: IPR001810; -.				
DR	PFAM: PF00400; WD40; 7.				
DR	PFAM: PF00646; F-box; 1.				
DR	PRINTS: PR00320; GPROTEINRPT.				
DR	PROSITE: PS00678; WD_REPEATS; UNKNOWN_6.				
KW	WD repeat.				
SQ	SEQUENCE 569 AA: 65049 MW: 6E7C4FE7BB520BA CRC64:				
Query Match	100.0%;	Score 3034;	DB 4;	Length 569;	
Best Local Similarity	100.0%;	Pred. No. 2.7e-244;			

Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDPAEAVLOEKALFKFMSSEEDCNNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLA 60
Db 1 MDPAEAVLOEKALFKFMSSEEDCNNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLA 60
QY 61 STAKTEVCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFEQWSESDOVEVEHL 120
Db 61 STAKTEVCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFEQWSESDOVEVEHL 120
QY 121 ISOMCHYOHGHINSYKPMLOQDFITLALPARGLDHAENILSYLDAKSLCAAEVCKEWY 180
Db 121 ISOMCHYOHGHINSYKPMLOQDFITLALPARGLDHAENILSYLDAKSLCAAEVCKEWY 180
QY 181 RYVSDGMLMKKLIERNVRTDSIMRGLAERGMGOYLFKNKPPDGNAPNSFYRALYPKII 240
Db 181 RYVSDGMLMKKLIERNVRTDSIMRGLAERGMGOYLFKNKPPDGNAPNSFYRALYPKII 240
QY 241 QDIETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
Db 241 QDIETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
QY 301 RILGHTGVSVCLOYDERVITITGSSDSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGM 360
Db 301 RILGHTGVSVCLOYDERVITITGSSDSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGM 360
QY 361 VTCSKDRSIAMWDMASPTDITLRVLGHRAAVNVDFDDKITYSAGDRTIKVWNTSTC 420
Db 361 VTCSKDRSIAMWDMASPTDITLRVLGHRAAVNVDFDDKITYSAGDRTIKVWNTSTC 420
QY 421 EFVFTLNGHKRGIACTLOYRDLRVVSSSDNTIRLMDIEGACLRVLEGHELVRCIFEDN 480
Db 421 EFVFTLNGHKRGIACTLOYRDLRVVSSSDNTIRLMDIEGACLRVLEGHELVRCIFEDN 480
QY 481 KRIYSAGYDGIKRWMDLVALDPAAPAGTCLRTLVHSGRVRLQDFEQIYSSSHDT 540
Db 481 KRIYSAGYDGIKRWMDLVALDPAAPAGTCLRTLVHSGRVRLQDFEQIYSSSHDT 540
QY 541 ILIMDFLNDPAQAEPSPSPRTTYTISR 569
Db 541 ILIMDFLNDPAQAEPSPSPRTTYTISR 569

RESULT 2
QY9297 PRELIMINARY: PRT: 605 AA.
AC QY9297:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE B-TRCP VARIANT E3RS-IKAPPAB (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=99075339; Pubmed=9859996;
RA Yaron A., Hatzudai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;
RT "Identification of the receptor component of the IkappaBalpha-
RT ubiquitin ligase."
RL Nature 396:590-594(1998).
DR EMBL: AF101784; AAD08702.1; -.
DR INTERPRO: IPR001680; -.
DR INTERPRO: IPR001810; -.
DR PAM: PF00440; WD40; 7.
DR PFAM: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_6.
FT NON_TER 1 1
SQ SEQUENCE 605 AA; 6886 MM; 4C67F3B7E400FD37 CRC64;

```

Query Match 99.1%; Score 3006; DB 4; Length 605;
 Best Local Similarity 94.0%; Pred. No. 6,3e-242;
 Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

```

QY 1 MDPAEAVLOEKALFKF-----NSSEEDC 24
Db 1 MDPAEAVLOEKALFKFMSMPRLMIGCSSLADSMPSRLCLVNPGTGALTARONSSEEDC 60
QY 25 NNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLASAMTENCVAKTKLANGTSSMIV 84
Db 25 NNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLASAMTENCVAKTKLANGTSSMIV 84
QY 61 NNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLASAMTENCVAKTKLANGTSSMIV 120
Db 61 NNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLASAMTENCVAKTKLANGTSSMIV 120
QY 85 PKORKLSASYEKEKELCVKFEQWSESDOVEVEHLISOMCHYOHGHINSYKPMLOQDF 144
Db 85 PKORKLSASYEKEKELCVKFEQWSESDOVEVEHLISOMCHYOHGHINSYKPMLOQDF 144
QY 121 PKORKLSASYEKEKELCVKFEQWSESDOVEVEHLISOMCHYOHGHINSYKPMLOQDF 180
Db 121 PKORKLSASYEKEKELCVKFEQWSESDOVEVEHLISOMCHYOHGHINSYKPMLOQDF 180
QY 145 ITALPARGLDHAENILSYLDAKSLCAAEVCKEWYRYVSDGMLMKKLIERNVRTDSLNR 204
Db 145 ITALPARGLDHAENILSYLDAKSLCAAEVCKEWYRYVSDGMLMKKLIERNVRTDSLNR 204
QY 181 ITALPARGLDHAENILSYLDAKSLCAAEVCKEWYRYVSDGMLMKKLIERNVRTDSLNR 240
Db 181 ITALPARGLDHAENILSYLDAKSLCAAEVCKEWYRYVSDGMLMKKLIERNVRTDSLNR 240
QY 205 GLAERGMGOYLFKNKPPDGNAPNSFYRALYPKIIDDIETIESNMRCGRHSLORIHCRS 264
Db 205 GLAERGMGOYLFKNKPPDGNAPNSFYRALYPKIIDDIETIESNMRCGRHSLORIHCRS 264
QY 241 GLAERGMGOYLFKNKPPDGNAPNSFYRALYPKIIDDIETIESNMRCGRHSLORIHCRS 300
Db 241 GLAERGMGOYLFKNKPPDGNAPNSFYRALYPKIIDDIETIESNMRCGRHSLORIHCRS 300
QY 265 ETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILTGHTGVSVCLOYDERVIYTS 324
Db 265 ETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILTGHTGVSVCLOYDERVIYTS 324
QY 301 ETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILTGHTGVSVCLOYDERVIYTS 360
Db 301 ETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILTGHTGVSVCLOYDERVIYTS 360
QY 325 SSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMVTCSKDRSIAMWDMASPTDITLRR 384
Db 325 SSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMVTCSKDRSIAMWDMASPTDITLRR 384
QY 361 SSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMVTCSKDRSIAMWDMASPTDITLRR 420
Db 361 SSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMVTCSKDRSIAMWDMASPTDITLRR 420
QY 385 VLVGHRAAVNVDFDDKITYSAGDRTIKVWNTSCFVFTLNGHKRGIACTLOYRDLRV 444
Db 385 VLVGHRAAVNVDFDDKITYSAGDRTIKVWNTSCFVFTLNGHKRGIACTLOYRDLRV 444
QY 421 VLVGHRAAVNVDFDDKITYSAGDRTIKVWNTSCFVFTLNGHKRGIACTLOYRDLRV 480
Db 421 VLVGHRAAVNVDFDDKITYSAGDRTIKVWNTSCFVFTLNGHKRGIACTLOYRDLRV 480
QY 445 SSSDNTIRLMDIEGACLRVLEGHELVRCIFEDNKRIYSGAYDGKIKWMDLVAAADPR 504
Db 445 SSSDNTIRLMDIEGACLRVLEGHELVRCIFEDNKRIYSGAYDGKIKWMDLVAAADPR 504
QY 481 SSSDNTIRLMDIEGACLRVLEGHELVRCIFEDNKRIYSGAYDGKIKWMDLVAAADPR 540
Db 481 SSSDNTIRLMDIEGACLRVLEGHELVRCIFEDNKRIYSGAYDGKIKWMDLVAAADPR 540
QY 505 APAGTCLRTLVHSGRVRLQDFEQIYSSSHDTILIMDFLNDPAQAEPSPSPRTY 564
Db 505 APAGTCLRTLVHSGRVRLQDFEQIYSSSHDTILIMDFLNDPAQAEPSPSPRTY 564
QY 541 APAGTCLRTLVHSGRVRLQDFEQIYSSSHDTILIMDFLNDPAQAEPSPSPRTY 600
Db 541 APAGTCLRTLVHSGRVRLQDFEQIYSSSHDTILIMDFLNDPAQAEPSPSPRTY 600

RESULT 3
QY9015 PRELIMINARY: PRT: 569 AA.
AC QY9015:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE B-TRCP VARIANT E3RS-IKAPPAB (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=99199275; Pubmed=10097128;
RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA Hatori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,
RA Nakayama K.-I.;
RT "Ubiquitin ligase Skp1/Cul1/F-box protein FMD1."
RT "Ubiquitin ligase Skp1/Cul1/F-box protein FMD1."
RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
FT NON_TER 1 1
SQ SEQUENCE FROM N.A.

```

RX MEDLINE-99075339; PubMed-9859996;
 RA Yaron A., Hatzubal A., Davis M., Lavon I., Amit S., Manning A.M.,
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;
 RT "Identification of the receptor component of the IkappaBalpha-
 RT ubiquitin ligase";
 RL Nature 396:590-594(1998)
 DR EMBL: AF081887; AAD17755.1; -;
 DR EMBL: AF099932; AAD08701.1; -;
 DR INTERPRO: IPR001680; -;
 DR INTERPRO: IPR001810; -;
 DR PFAM: PF00400; WD40; 7;
 DR PFAM: PF00646; F-box; 1;
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_6.
 KW Ligase.
 SQ SEQUENCE 569 AA; 65105 MW; BC7D6544815B2296 CRC64;

Query Match 98.8%; Score 2997; DB 11; Length 569;
 Best Local Similarity 98.6%; Pred. No. 3.2e-241;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNSSREDCNNGEPPRKTIIPKNSLRQTYNSCARCLNOETVCLA 60
 DB 1 MDPAEAVLOEKALKFNSSREDCNNGEPPRKTIIPKNSLRQTYNSCARCLNOETVCLT 60
 QY 61 STAMKTENCYAKTKLANGTSSMTVPKOKLSASYEKEKELCVYFPQWSSDQVEVEHL 120
 DB 61 STAMKTENCYAKTKLANGTSSMTVPKOKLSASYEKEKELCVYFPQWSSDQVEVEHL 120
 QY 121 ISOMCHYOHGHINSYLYKPMLOQDFITALPARGLDHAENILSYLDKSLCAELVCKEWY 180
 DB 121 ISOMCHYOHGHINSYLYKPMLOQDFITALPARGLDHAENILSYLDKSLCAELVCKEWY 180
 QY 181 RYVSDGMLMKKLLERAVRTDLSMRGLAERGMQYLFKNKPPDGNAPNSFYALPKII 240
 DB 181 RYVSDGMLMKKLLERAVRTDLSMRGLAERGMQYLFKNKPPDGNAPNSFYALPKII 240
 QY 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 DB 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 QY 301 RILTGHTGSLVLCQYDERVYITGSSDSTYRVWVDNTEGEMINTLIHCEAVLHLRFNNGM 360
 DB 301 RILTGHTGSLVLCQYDERVYITGSSDSTYRVWVDNTEGEMINTLIHCEAVLHLRFNNGM 360
 QY 361 VTCSKRSIAVWDMASPTDITLRVLYGHRAAVNVDFDQKIVYSAGDRTIKVMNTSTC 420
 DB 361 VTCSKRSIAVWDMASPTDITLRVLYGHRAAVNVDFDQKIVYSAGDRTIKVMNTSTC 420
 QY 421 EFVRTLNGHKGRIACQYRDRLVYVSSSDNTIRLMDIEGACLRVLEGHELVRCIRFDN 480
 DB 421 EFVRTLNGHKGRIACQYRDRLVYVSSSDNTIRLMDIEGACLRVLEGHELVRCIRFDN 480
 QY 481 KRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVEHSGRVERLQDFEQIVSSSHDDT 540
 DB 481 KRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVEHSGRVERLQDFEQIVSSSHDDT 540
 QY 541 ILIMDFLNDPAQAEPSPRSRTTYISR 569
 DB 541 ILIMDFLNDPAQAEPSPRSRTTYISR 569

RESULT 4
 092159
 ID 092159 PRELIMINARY; PRT: 569 AA.
 AC 092159;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-JUN-2000 (TREMBLrel. 14, last annotation update)
 DE BETA-TRANSDUCIN REPEAT CONTAINING PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RX (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99145465; PubMed-9990853;
 RA Spencer E., Jiang J., Chen Z.J.;
 RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
 RT S1mb/Beta-Trcp";
 RL Genes Dev. 13:284-294(1999).
 DR EMBL: AF112979; AAD04181.1; -;
 DR INTERPRO: IPR001680; -;
 DR INTERPRO: IPR001810; -;
 DR PFAM: PF00400; WD40; 7;
 DR PFAM: PF00646; F-box; 1;
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_6.
 SQ SEQUENCE 569 AA; 65047 MW; BC7CA4A4815BED96 CRC64;

Query Match 98.5%; Score 2990; DB 11; Length 569;
 Best Local Similarity 98.4%; Pred. No. 1.2e-240;
 Matches 560; Conservative 3; Mismatches .6; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNSSREDCNNGEPPRKTIIPKNSLRQTYNSCARCLNOETVCLA 60
 DB 1 MDPAEAVLOEKALKFNSSREDCNNGEPPRKTIIPKNSLRQTYNSCARCLNOETVCLT 60
 QY 61 STAMKTENCYAKTKLANGTSSMTVPKOKLSASYEKEKELCVYFPQWSSDQVEVEHL 120
 DB 61 STAMKTENCYAKTKLANGTSSMTVPKOKLSASYEKEKELCVYFPQWSSDQVEVEHL 120
 QY 121 ISOMCHYOHGHINSYLYKPMLOQDFITALPARGLDHAENILSYLDKSLCAELVCKEWY 180
 DB 121 ISOMCHYOHGHINSYLYKPMLOQDFITALPARGLDHAENILSYLDKSLCAELVCKEWY 180
 QY 181 RYVSDGMLMKKLLERAVRTDLSMRGLAERGMQYLFKNKPPDGNAPNSFYALPKII 240
 DB 181 RYVSDGMLMKKLLERAVRTDLSMRGLAERGMQYLFKNKPPDGNAPNSFYALPKII 240
 QY 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 DB 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 QY 301 RILTGHTGSLVLCQYDERVYITGSSDSTYRVWVDNTEGEMINTLIHCEAVLHLRFNNGM 360
 DB 301 RILTGHTGSLVLCQYGERVYITGSSDSTYRVWVDNTEGEMINTLIHCEAVLHLRFNNGM 360
 QY 361 VTCSKRSIAVWDMASPTDITLRVLYGHRAAVNVDFDQKIVYSAGDRTIKVMNTSTC 420
 DB 361 VTCSKRSIAVWDMASPTDITLRVLYGHRAAVNVDFDQKIVYSAGDRTIKVMNTSTC 420
 QY 421 EFVRTLNGHKGRIACQYRDRLVYVSSSDNTIRLMDIEGACLRVLEGHELVRCIRFDN 480
 DB 421 EFVRTLNGHKGRIACQYRDRLVYVSSSDNTIRLMDIEGACLRVLEGHELVRCIRFDN 480
 QY 481 KRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVEHSGRVERLQDFEQIVSSSHDDT 540
 DB 481 KRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVEHSGRVERLQDFEQIVSSSHDDT 540
 QY 541 ILIMDFLNDPAQAEPSPRSRTTYISR 569
 DB 541 ILIMDFLNDPAQAEPSPRSRTTYISR 569

RESULT 5
 09R1G7
 ID 09R1G7 PRELIMINARY; PRT: 569 AA.
 AC 09R1G7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBLrel. 14, last annotation update)
 DE BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Winston J., Elledge S.J., Harper J.W.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF110396; AAD41025.1; -
 DR INTERPRO; IPR001680; -
 DR INTERPRO; IPR001810; -
 DR PFAM; PF004400; WD40; 7
 DR PFAM; PF00646; F-box; 1
 DR PRINTS; PR00320; GPROTEINRPT.
 DR PROSITE; PS00678; WD_REPEATS;
 SQ SEQUENCE 569 AA; 65209 MM; EGGDCAD28D51D9D CRC64;

Query Match 98.2%; Score 2978; DB 11; Length 569;
 Best Local Similarity 97.9%; Pred. No. 1.2e-239;
 Matches 557; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MDPAAVLOEKALFKPMSEREDCNGEPKRIPEKNSLRQYNSCARLCTNOETVCLA 60
 DB 1 MDPAAVLOEKALFKPMSEREDCNGEPKRIPEKNSLRQYNSCARLCTNOETVCLT 60
 QY 61 STAKTEKNCVAKTKLANGSSMIVPKQKLSASYEKEKELCVKFEQMSDQVEFEVHL 120
 DB 61 STAKTEKNCVAKTKLANGSSMIVPKQKLSASYEKEKELCVKFEQMSDQVEFEVHL 120
 QY 121 ISOMCHYOHGHINSYLPMLORDFTALPARGLDHAENILSYLDAKSLCAELVCKEY 180
 DB 121 ISOMCHYOHGHINSYLPMLORDFTALPARGLDHAENILSYLDAKSLCAELVCKEY 180
 QY 181 RYTSQGMKLLIERVMTDLSMRGLAERGGQYLFKNRPPDGNAPNSFYALYKII 240
 DB 181 RYTSQGMKLLIERVMTDLSMRGLAERGGQYLFKNRPPDGNAPNSFYALYKII 240
 QY 241 QDIETIESNMRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLRDNTIKIMDKTECK 300
 DB 241 QDIETIESNMRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLRDNTIKIMDKTECK 300
 QY 301 RILGHTGSVYCLQYDERVIITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
 DB 301 RILGHTGSVYCLQYDERVIITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
 QY 361 VTCSKDRSIAVMDMSPDTITLRVLYGHRAAVNVDFDQKIYASAGDRTIKVNTSTC 420
 DB 361 VTCSKDRSIAVMDMSPDTITLRVLYGHRAAVNVDFDQKIYASAGDRTIKVNTSTC 420
 QY 421 EFVRLNGHKGRIACLOYRDLVYSSGSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
 DB 421 EFVRLNGHKGRIACLOYRDLVYSSGSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
 QY 481 KRIYSGAVDGKIKVMDLVAAIDPPAPAGTLCRLTVLHSGAVFRLQDFEQIVSSHDDT 540
 DB 481 KRIYSGAVDGKIKVMDLVAAIDPPAPAGTLCRLTVLHSGAVFRLQDFEQIVSSHDDT 540
 QY 541 ILIMDFLNDPAAQAEPSPSRITYISR 569
 DB 541 ILIMDFLNDPAAQAEPSPSRITYISR 569

RESULT 6
 Q9P2S8 PRELIMINARY; PRT; 529 AA.
 AC Q9P2S8; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2 ISOFORM B.
 GN BTRCP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-FETAL LUNG;
 RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,
 RA Katoh M.;
 RT "Molecular cloning and genomic structure of the betaTRCP2 gene on
 RT chromosome 5q35.1."
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).
 DR EMBL; AB033280; BAA92330.1; -
 SQ SEQUENCE 529 AA; 60897 MM; 53D181B789313D0E CRC64;

Query Match 79.1%; Score 2400; DB 4; Length 529;
 Best Local Similarity 79.3%; Pred. No. 1.5e-191;
 Matches 451; Conservative 43; Mismatches 35; Indels 40; Gaps 4;

QY 1 MDPAAVLOEKALFKPMSEREDCNGEPKRIPEKNSLRQYNSCARLCTNOETVCLA 60
 DB 1 MDPAAVLOEKALFKPMSEREDCNGEPKRIPEKNSLRQYNSCARLCTNOETVCLT 60
 QY 61 STAKTEKNCVAKTKLANGSSMIVPKQKLSASYEKEKELCVKFEQMSDQVEFEVHL 120
 DB 61 STAKTEKNCVAKTKLANGSSMIVPKQKLSASYEKEKELCVKFEQMSDQVEFEVHL 120
 QY 37- -----ISNGTSVIVSRKRRESBGNQYKEDKLCIKYDQMSDQVEFEVHL 82
 DB 37- -----ISNGTSVIVSRKRRESBGNQYKEDKLCIKYDQMSDQVEFEVHL 82
 QY 121 ISOMCHYOHGHINSYLPMLORDFTALPARGLDHAENILSYLDAKSLCAELVCKEY 180
 DB 121 ISOMCHYOHGHINSYLPMLORDFTALPARGLDHAENILSYLDAKSLCAELVCKEY 180
 QY 83 ISMCHYOHGHINSYLPMLORDFTALPARGLDHAENILSYLDAKSLCAELVCKEY 142
 DB 83 ISMCHYOHGHINSYLPMLORDFTALPARGLDHAENILSYLDAKSLCAELVCKEY 142
 QY 181 RYTSQGMKLLIERVMTDLSMRGLAERGGQYLFKNRPPDGNAPNSFYALYKII 240
 DB 181 RYTSQGMKLLIERVMTDLSMRGLAERGGQYLFKNRPPDGNAPNSFYALYKII 240
 QY 143 RYTSQGMKLLIERVMTDLSMRGLAERGGQYLFKNRPPDGNAPNSFYALYKII 200
 DB 143 RYTSQGMKLLIERVMTDLSMRGLAERGGQYLFKNRPPDGNAPNSFYALYKII 200
 QY 241 QDIETIESNMRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLRDNTIKIMDKTECK 300
 DB 241 QDIETIESNMRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLRDNTIKIMDKTECK 300
 QY 201 QDIETIESNMRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLRDNTIKIMDKTECK 260
 DB 201 QDIETIESNMRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLRDNTIKIMDKTECK 260
 QY 301 RILGHTGSVYCLQYDERVIITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
 DB 301 RILGHTGSVYCLQYDERVIITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360
 QY 261 KVLGHTGSVYCLQYDERVIITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 320
 DB 261 KVLGHTGSVYCLQYDERVIITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 320
 QY 361 VTCSKDRSIAVMDMSPDTITLRVLYGHRAAVNVDFDQKIYASAGDRTIKVNTSTC 420
 DB 361 VTCSKDRSIAVMDMSPDTITLRVLYGHRAAVNVDFDQKIYASAGDRTIKVNTSTC 420
 QY 421 EFVRLNGHKGRIACLOYRDLVYSSGSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
 DB 421 EFVRLNGHKGRIACLOYRDLVYSSGSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
 QY 481 KRIYSGAVDGKIKVMDLVAAIDPPAPAGTLCRLTVLHSGAVFRLQDFEQIVSSHDDT 540
 DB 481 KRIYSGAVDGKIKVMDLVAAIDPPAPAGTLCRLTVLHSGAVFRLQDFEQIVSSHDDT 540
 QY 541 ILIMDFLNDPAAQAEPSPSRITYISR 569
 DB 501 ILIMDFLNDPAAQAEPSPSRITYISR 529

RESULT 7
 Q9UKB1 PRELIMINARY; PRT; 542 AA.
 AC Q9UKB1; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE F-BOX PROTEIN FBW1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2 ISOFORM C).
 GN FBW1B OR BTRCP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Geniatarelli C., Chiau D.S., Guardavaccaro D., Parks W., Vidal M.,
 Pagano M.;
 RT "Identification of a family of human F-box proteins.";
 RL Curr. Biol. 9:1177-1179(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FETAL LUNG;
 RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,
 Katoh M.;
 RT "Molecular cloning and genomic structure of the betatrop2 gene on
 chromosome 5q35.1";
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).
 DR EMBL: AF176022; AAF04528.1; -;
 DR EMBL: AB033281; BAA92331.1; -;
 DR INTERPRO: IPR001680; -;
 DR INTERPRO: IPR001810; -;
 DR PFM: PF00400; WD40; 7;
 DR PFM: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR PROSITE: PS00678; WD_REPEATS;
 SO SEQUENCE 542 AA; 62090 MW; 7CD40087EFA5C8A CRC64;

Query Match 78.6%; Score 2384.5; DB 4; Length 542;
 Best Local Similarity 79.0%; Pred. No. 2.9e-190;
 Matches 451; Conservative 48; Mismatches 41; Indels 31; Gaps 7;

QY 1 MDPAAVLOEAKLAFNNSSRECCNNGEPKRIIPEKNSIRQYINSGARCLNOEYCLA 60
 DB 1 MEP-DSVIEDKTELKCS-----VP-----RSIMLGCANLV---ESMCAL 36
 QY 61 S--TAKTKENCVAKTKLANGTSSMIYVPOKRLSAYEKEKELCVKYPEQSESDQVEFE 118
 DB 37 SCLOSPSVACL---QISNCTSSVIYSKRKPSGKNGVCKEKLCKIKTFDQSESDQVEFE 93
 QY 119 HLISQCHVOHGINSYKLMQORDFTTALPARGLDHIAENILSYLDAKSIACAELVCKE 178
 DB 94 HLISRMCHVOHGINSYKLMQORDFTTALPARGLDHIAENILSYLDAKSIACAELVCKE 153
 QY 179 WYVYSDGMIMKLLIERMVTDSLMRGLAERKRGQYLFKNKPPDGNAPNSFYRALYK 238
 DB 154 WQVVISSEGMIMKLLIERMVTDSLMRGLAERKRGQYLFKNKPPDGNAPNSFYRALYK 211
 QY 239 IIDDIETIESNMRCGRHSIORICRSETSKGYVCLQYDDQKIYSGLRDNTIKIMDKTLE 298
 DB 212 IIDDIETIESNMRCGRHSIORICRSETSKGYVCLQYDDQKIYSGLRDNTIKIMDKTLE 271
 QY 299 CKRILGHTGSVCLQYDERVIITGSSDSYVRYWDVTGEMLTLLIHCAVLAHLRFNNG 358
 DB 272 CLKVLGHTGSVCLQYDERVIITGSSDSYVRYWDVTGEMLTLLIHCAVLAHLRFNNG 331
 QY 359 MMYTCSKDRSIAVMDASPTDITLRVLYVGHRAAVNVVDEDDKIYVSASGDRITKYWNTS 418
 DB 332 LMVTCSDKRSIAVMDASPTDITLRVLYVGHRAAVNVVDEDDKIYVSASGDRITKYWNTS 391
 QY 419 TCEFVRLNKHKGKIGACLOYRDLRYVSGSSDNTIRLMDIECGACRLVLEHHEELVRCIRF 478
 DB 392 TCEFVRLNKHKGKIGACLOYRDLRYVSGSSDNTIRLMDIECGACRLVLEHHEELVRCIRF 451
 QY 479 DNKRIYSGAIDGKIKYWDVAALDPRAPAGTLCIRTLVEHSGVFRLOPFEQIVSSSD 538
 DB 452 DNKRIYSGAIDGKIKYWDVAALDPRAPAGTLCIRTLVEHSGVFRLOPFEQIVSSSD 511
 QY 539 DTLLIMDFLNDPAAQAEPPRSPRTYTYISR 569
 DB 512 DTLLIMDFLNDPAAQAEPPRSPRTYTYISR 542

RESULT 8
 Q9Y4C6 PRELIMINARY; PRT; 550 AA.
 ID Q9Y4C6

AC Q9Y4C6;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, last annotation update)
 DE KIA00696 PROTEIN (FRAGMENT).
 GN KIA00696.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=98403880; PubMed=9734811;
 RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
 Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. X.
 The complete sequences of 100 new cDNA clones from brain which can
 code for large proteins in vitro.";
 RL DNA Res. 5:169-176(1998).
 DR EMBL: AB014596; BAA31671.1; -;
 DR INTERPRO: IPR001680; -;
 DR INTERPRO: IPR001810; -;
 DR PFM: PF00400; WD40; 7;
 DR PFM: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR PROSITE: PS00678; WD_REPEATS;
 FT NON_TER 1
 SO SEQUENCE 550 AA; 62702 MW; FDE42B5BF3D19C8A CRC64;

Query Match 78.6%; Score 2384.5; DB 4; Length 550;
 Best Local Similarity 79.0%; Pred. No. 3e-190;
 Matches 451; Conservative 48; Mismatches 41; Indels 31; Gaps 7;

QY 1 MDPAAVLOEAKLAFNNSSRECCNNGEPKRIIPEKNSIRQYINSGARCLNOEYCLA 60
 DB 9 MEP-DSVIEDKTELKCS-----VP-----RSIMLGCANLV---ESMCAL 44
 QY 61 S--TAKTKENCVAKTKLANGTSSMIYVPOKRLSAYEKEKELCVKYPEQSESDQVEFE 118
 DB 45 SCLOSPSVACL---QISNCTSSVIYSKRKPSGKNGVCKEKLCKIKTFDQSESDQVEFE 101
 QY 119 HLISQCHVOHGINSYKLMQORDFTTALPARGLDHIAENILSYLDAKSIACAELVCKE 178
 DB 102 HLISRMCHVOHGINSYKLMQORDFTTALPARGLDHIAENILSYLDAKSIACAELVCKE 161
 QY 179 WYVYSDGMIMKLLIERMVTDSLMRGLAERKRGQYLFKNKPPDGNAPNSFYRALYK 238
 DB 162 WQVVISSEGMIMKLLIERMVTDSLMRGLAERKRGQYLFKNKPPDGNAPNSFYRALYK 219
 QY 239 IIDDIETIESNMRCGRHSIORICRSETSKGYVCLQYDDQKIYSGLRDNTIKIMDKTLE 298
 DB 220 IIDDIETIESNMRCGRHSIORICRSETSKGYVCLQYDDQKIYSGLRDNTIKIMDKTLE 279
 QY 299 CKRILGHTGSVCLQYDERVIITGSSDSYVRYWDVTGEMLTLLIHCAVLAHLRFNNG 358
 DB 280 CLKVLGHTGSVCLQYDERVIITGSSDSYVRYWDVTGEMLTLLIHCAVLAHLRFNNG 339
 QY 359 MMYTCSKDRSIAVMDASPTDITLRVLYVGHRAAVNVVDEDDKIYVSASGDRITKYWNTS 418
 DB 340 LMVTCSDKRSIAVMDASPTDITLRVLYVGHRAAVNVVDEDDKIYVSASGDRITKYWNTS 399
 QY 419 TCEFVRLNKHKGKIGACLOYRDLRYVSGSSDNTIRLMDIECGACRLVLEHHEELVRCIRF 478
 DB 400 TCEFVRLNKHKGKIGACLOYRDLRYVSGSSDNTIRLMDIECGACRLVLEHHEELVRCIRF 459
 QY 479 DNKRIYSGAIDGKIKYWDVAALDPRAPAGTLCIRTLVEHSGVFRLOPFEQIVSSSD 538
 DB 460 DNKRIYSGAIDGKIKYWDVAALDPRAPAGTLCIRTLVEHSGVFRLOPFEQIVSSSD 519
 QY 539 DTLLIMDFLNDPAAQAEPPRSPRTYTYISR 569
 DB 512 DTLLIMDFLNDPAAQAEPPRSPRTYTYISR 542

Db 520 DTLLIMDFLNPSPSAONETSPSRRTYYSR 550

RESULT 9

Q9P2S9 PRELIMINARY; PRT; 508 AA.

AC Q9P2S9; 01-JUN-1998 (TRENBLREL. 15, Created)

DT 01-OCT-2000 (TRENBLREL. 15, last sequence update)

DE F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2 ISOFORM A.

GN BTRCP2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=FETAL LUNG;

RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M., Katoh M.;

RT "Molecular cloning and genomic structure of the betaTRCP2 gene on chromosome 5q35.1."

RL Biochem. Biophys. Res. Commun. 269:103-109(2000).

DR EMBL: AB033279; BAA92329.1; -

SO SEQUENCE 508 AA; 58422 MW; 05C691DB32A91B50 CRC64;

Query Match 77.7%; Score 2358; DB 4; Length 508;

Best Local Similarity 86.7%; Pred. No. 4.3e-188;

Matches 436; Conservative 34; Mismatches 31; Indels 2; Gaps 1;

QY 67 ENCVAKTRKLANGTSMIVPQKRLASAEKKEKELCVKFFQMSDDOFEVEHLISOMCH 126

Db 8 EDKTELMSIGTSYVSRKRPSEGNQKCKDKCYFDQMSDDOFEVEHLISOMCH 67

QY 127 YOHGHINSYLPMLQRFETALPARGLDHAENILSYLDKSLCAELVCKEWMRYTSDG 186

Db 68 YOHGHINSYLPMLQRFETALPARGLDHAENILSYLDKSLCAELVCKEWMRYTSDG 127

QY 187 MLMKKLIERMVRTSLWRGLAERNGOYLFFKNRPPGNAPNSFYALYPKIIQDIETI 246

Db 128 MLMKKLIERMVRTSLWRGLAERNGOYLFFKNRPPGNAPNSFYALYPKIIQDIETI 185

QY 247 ESNMRGHSRHSIORHCRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKTECKRIITGH 306

Db 186 ESNMRGHSRHSIORHCRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKTECKRIITGH 245

QY 307 TGSVLCLOYDERVIITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMAYTCKSD 366

Db 246 TGSVLCLOYDERVIITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMAYTCKSD 305

QY 367 RSIWVMDASTDTILTRVLYGHRAAVNVDFDCKYIVSASGDTIKVMNTSCFEFRTL 426

Db 306 RSIWVMDASTDTILTRVLYGHRAAVNVDFDCKYIVSASGDTIKVMNTSCFEFRTL 365

QY 427 NGHRGIACTLOYRDRLVYSGSSDNTIRLMDIEGACLRVLEGHELYRCIRFDMKRIYVG 486

Db 366 NGHRGIACTLOYRDRLVYSGSSDNTIRLMDIEGACLRVLEGHELYRCIRFDMKRIYVG 425

QY 487 AYDGKIKWDLVAALDPRAPAGTLCRLTVLHSGRVRLQDFEQIYSSSHDITLIMDF 546

Db 426 AYDGKIKWDLVAALDPRAPAGTLCRLTVLHSGRVRLQDFEQIYSSSHDITLIMDF 485

QY 547 LNDPAQAQEPSPSRRTYYSR 569

Db 486 LNDPAQAQEPSPSRRTYYSR 508

RESULT 10

ID 044382 PRELIMINARY; PRT; 510 AA.

AC 044382;

DT 01-JUN-1998 (TRENBLREL. 06, Created)

DT 01-JUN-1998 (TRENBLREL. 06, last sequence update)

DT 01-OCT-2000 (TRENBLREL. 15, last annotation update)

DE SLIMB.

GN SLIMB.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=98121115; PubMed=9461217;

RA Jiang J., Struhl G.;

RT "Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb.";

RL Nature 391:493-496(1998).

DR EMBL: AF032878; AAC38852.1; -

DR FLYBASE: FBgn0023423; slmb.

DR INTERPRO: IPR001680; -

DR INTERPRO: IPR001610; -

DR PFAM: PF00400; WD40; 7.

DR PFAM: PF00646; F-box; 1.

DR PRINTS: PR00320; GPROTEINRPT.

DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_5.

KW Repeat.

SO SEQUENCE 510 AA; 58952 MW; DB0243D3730A5E8 CRC64;

Query Match 68.9%; Score 2089.5; DB 5; Length 510;

Best Local Similarity 78.3%; Pred. No. 9.8e-166;

Matches 394; Conservative 48; Mismatches 54; Indels 7; Gaps 4;

QY 64 MKTENVAATKLANG---SSMIVPQKRLAS--YEKEKELCVKFFQMSDDOFEVE 118

Db 4 METDKIMDET-N-SNAQFTTMYLDPVRKRDSSPTTYTERELCFYFQWSSGGVDVE 62

QY 119 HLTSOMCHYOHGHINSYLPMLQRFETALPARGLDHAENILSYLDKSLCAELVCKE 178

Db 63 HLTSOMCHYOHGHINSYLPMLQRFETALPARGLDHAENILSYLDKSLCAELVCKE 122

QY 179 WRTVSDGLMLKKLIERMVRTSLWRGLAERNGOYLFFKNRPPGNAPNSFYALYPK 238

Db 123 WRTVSDGLMLKKLIERMVRTSLWRGLAERNGOYLFFKNRPPGNAPNSFYALYPK 181

QY 239 IIODITFESNMRCGRHSIORHCRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKTE 298

Db 182 IIODITFESNMRCGRHSIORHCRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKTE 241

QY 299 CKRILTGHGTVCLQYDERVIITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNG 358

Db 242 CKRILTGHGTVCLQYDERVIITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNG 301

QY 359 MMYTCKSDRSIAWMDASTDTILTRVLYGHRAAVNVDFDCKYIVSASGDTIKVMNTS 418

Db 302 MMYTCKSDRSIAWMDASTDTILTRVLYGHRAAVNVDFDCKYIVSASGDTIKVMNTS 361

QY 419 TCFEVRTLNGHRGIACTLOYRDRLVYSGSSDNTIRLMDIEGACLRVLEGHELYRCIRF 478

Db 362 TCFEVRTLNGHRGIACTLOYRDRLVYSGSSDNTIRLMDIEGACLRVLEGHELYRCIRF 421

QY 479 DNKRIYSGAYDGKIKWDLVAALDPRAPAGTLCRLTVLHSGRVRLQDFEQIYSSSHD 538

Db 422 DNKRIYSGAYDGKIKWDLVAALDPRAPAGTLCRLTVLHSGRVRLQDFEQIYSSSHD 481

QY 539 DTLLIMDFLNDPAQAQEPSPSR 561

Db 482 DTLLIMDFLNDPAQAQEPSPSR 504

RESULT 11

ID 09VDE3 PRELIMINARY; PRT; 510 AA.

AC 09VDE3


```

OY 123 OMCHXHGHNISLXPMJQROFETALBPARGDHAENILSYLDAKSJCAELVCKEMRY 182
Db 51 SLSEGVAVAYVNNHRSLLFTDFTEVP-----EYVSFLVSTYLOLDJCKCKLSKMRKL 106
OY 183 TSDGMLMKLL-----ERMYETDSLARG-----LAERGWG----- 213
Db 107 LEDGIMKALYMKQWVFNENVLNEFNAWRTHKFPQPRENFLOKONITGPGYTFPQ 166
OY 214 OYLFKRNKPPDONAPNSFYALPKIITQDLETTESNNCRSHLSQTRHCS----- 264
Db 167 QYIF-----DSNGRPLNMSTLY---KEAHLDSDNRHGRFLVSTFNPSIRFPADQF 217
OY 265 -ETSGKVYCLQYDQCKIVSGLDYTFIKIMDKNTELECKRIITGTHSYLQYDER--VII 321
Db 218 RATIDSVYCVQYDDEIMVSGSKDPTVSVMQVNSFLIKYLGHSYVLCJDFCRRNLV 277
OY 322 TGSSDSYRVWVDVTGEMKLTLLHNCAYVLHLRFNNCMATYCSKDSIAW--DMASPD 379
Db 278 SGSSDSYIIMDMONRRPLKYFGHTNVLGVVYSENYIISSRDHARWRIADATSPAE 337
OY 380 ITLRLVGLGHRAAVNVYDEDK--YIYASAGDPTIKWNTSTGEFRTLGHKRGJACQ 437
Db 338 ACM-HVLRGLHASVNSQYSKTLIYTAASDRILKRWDTTGHGCIIRIHANRGJACQ 396
OY 438 YRDLRVYSGSSDNTRIUMDIECGACLVLEGHEELVNCIRFDNKRIYSGAYDKIKWDL 497
Db 397 YNGKFVYSGSDELTRIFEAASSGKILMLQGHEDLRTVFENDEK IYSGGYDGTVRIMN- 455
OY 498 VAALDRAPRATLQRLVHE-----SGRVFLQDDEQOYIYSSSHDPTILLMDP 546
Db 456 -----FNTGEQCVLHNSRNSYVCLQDHRRIITACHTSSEIWNF 497

```

RESULT	13		
ID	09VZF4	PRELIMINARY;	PRT; 1326 AA.
AC	09VZF4		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)		
DE	CG15010 PROTEIN.		
GN	CG15010.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wotman J.R., Yandell M.D., Zhang O., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abiril J.F., Agbayanl A., An H.-J., Andrews-Plankkoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahle K., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Dowes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Gang N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jatelli M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		

RA Liu X.H., Mattetelli B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paleb J.M.,
 RA Palazzolo M., Petlan G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinelt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Splier E., Spradling A.C., Stapleton M., Strong K., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weltschok G.M., Weltschbach J.,
 RA Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C. ;
 RT "The genome sequence of *Drosophila melanogaster*." ;
 RL Science 287:2185-2195(2000) .
 DR EMBL: AE003480; AAF47869.1 ; -
 DR FLYBASE; FBgn0035516; CG15010.
 DR INTERPRO: IPR001680; -
 DR INTERPRO: IPR001810; -
 DR PFAM: PF004400; WD40; 7.
 DR PFAM: PF00646; F-box; 1.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PROSITE: PS00678; WD_REPEATS.5.
 SO SEQUENCE 1326 AA; 141360 MW; 3542C873CEFA3027F CRC64;

```

Query Match: 21.8%; Score 661; DB 5; Length 1326;
Best Local Similarity 34.9%; Pred. No. 3,1e-46;
Matches 158; Conservative 76; Mismatches 167; Indels 52; Gaps 12.

OY 105 FEQMSSESDQVEFHEHLISQWCHYOHGHNISYLKAPMLQDFETPALPARGLDHIAENILSYL 164
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
DB 852 FQNSHVERRLADRLDHODPQVNVHMMKVIYEPQGFDFISLPR---ELALFVLSYL 907
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
OY 165 DAKSLCAELVCEWYRVVTSQGLMKKLLIERNVVRTDSLMRGLAERMGQYLFKNKP-- 222
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
DB 908 EPKDLRAAQTCRSWRFCDNLLMKCKRKA-----QILAEP-----SDRKRG 953
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
OY 223 -DGNAP-NSFYVALPKTIIDETTESNWRGSRHSIORICRSEPTSG-----VYCLOY 275
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
DB 954 RDGMPPRIASPMVAATYR--QHI--TENMR-----SRPRKRYLKGHDHVIYTCLOP 1003
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
OY 276 DDOKIVSGLRDNTIKIMDKNTELECKRIILGHTGSVLCQYDERVITIGSSDSVYRWVDV 335
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
DB 1004 SGNRIYVSGSDNTILTKVMSAVNGKCLERTLVTGHGVWSSQMGNIILISGSDPTFLKVMDD 1063
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
OY 336 TGEMLNTLLIHCCAVYLHLRFNNCMAYTCKSDSIAMVPMASPTDITLRYLYGHRVAVN 395
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
DB 1064 SGACVHTLQHTSTVTKRMHLHGSKVYSSGSDRTFLKVMQIDEGSC--HYLVGLAAVRC 1120
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
OY 396 VDFPDKTVIASAGDRTIKVMNTSTCFEYVTLTGHRGLACIQDYDRLVYSSGSDNTIRLM 455
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
DB 1121 VQYGLKVLGAGVDYWKVKITHPRQDCLTTLQGHINRRYSLSQFQGLNVYSSSDTSTIRW 1180
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
OY 456 DIECGACRLVLEGHHELVRCIRFDRNKRIVSAGYADKIKVMDLVAALDPRAPAGTLCRTL 515
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
DB 1181 DVEITGNCKHTLMGHSITSGMELRQIILVSGNADSTVAKVMDITTG-----QCIQTL 1231
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
OY 516 V---EHSGRFRLQDFEFQIVSSSHDPTLLIWD 545
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::
DB 1232 SGPNNKHSAAVTCLOFNSRFVYVSSDGTIVKWD 1264
      +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +::: +:::

RESULT 14
Q9NUM6 PRELIMINARY; PRT; 553 AA.
AC Q9NUM6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, last annotation update)
DE CDNA FLJ1071 FIS, CLONE PLACE1004937, MODERATELY SIMILAR TO SEL-10
  PROTEIN.

```


THIS PAGE BLANK (USPTO)